

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**

91926

Access DB#

# SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: M. A. WALICKA Examiner #: 782201 Date: 04/21/03  
Art Unit: 1652 Phone Number 305-7270 Serial Number: 101050200  
Mail Box and Bldg/Room Location: 10D06 Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: Angiogenesis - novel - 2 Peptide Substrates  
Inventors (please provide full names): FOURIE et al and method

Earliest Priority Filing Date: 01/16/2002

\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Please search

SEQ ID NO: 8 } AA

SEQ ID NO: 1 } DNA

Thank you

Please RUSH

Thank you.

Malick

C. Chan  
Rush

Edward Hart  
Technical Info. Specialist  
STIC/Biotech  
CMI 6B02 Tel: 305-9203

## STAFF USE ONLY

	Type of Search	Vendors and cost where applicable
Searcher: _____	NA Sequence (#) <u>2</u> STN _____	
Searcher Phone #: _____	AA Sequence (#) <u>2</u> Dialog _____	
Searcher Location: _____	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: <u>4/21/03</u>	Bibliographic _____	Dr. Link _____
Date Completed: <u>4/26/03</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep & Review Time: _____	Fulltext _____	Sequence Systems <u>1H/02</u>
Clerical Prep Time: _____	Patent Family _____	WWW/Internet _____
Online Time: _____	Other _____	Other (specify) _____

BEST AVAILABLE COPY

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 14:43:36 ; Search time 20.4345 Seconds

(without alignments)  
708.413 Million cell updates/sec

Title: US-10-050-200-9

Perfect score: 2614

Sequence: 1 MLGMASTLLCAFRLLPLAAV.....LDLPRKQILGDKYKDDDDKG 492

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2549	97.5	930	4	US-09-122-126B-15
2	2077	79.5	930	4	US-09-369-364A-2
3	829.5	31.7	967	4	US-09-130-491-2
4	732.5	28.0	905	4	US-09-369-364A-9
5	658.5	25.2	837	4	US-09-122-126B-2
6	567.5	21.7	245	4	US-09-369-364A-11
7	565.5	21.6	1882	4	US-09-369-364A-13
8	560.5	21.4	874	4	US-09-369-364A-15
9	549.5	21.0	608	4	US-09-130-491-13
10	522.5	20.0	997	4	US-09-369-364A-7
11	462.5	17.7	1224	4	US-09-930-872-4
12	441.5	16.9	491	4	US-09-369-364A-5
13	413	15.8	859	4	US-09-369-364A-17
14	407	15.6	1205	4	US-09-491-522-11
15	388	14.8	1081	4	US-09-369-364A-1
16	387	14.8	566	4	US-09-491-522-7
17	374	14.3	1211	4	US-09-491-522-5
18	373	14.3	518	4	US-09-369-364A-22
19	313.5	12.0	802	4	US-09-632-098-2
20	313.5	12.0	812	4	US-09-632-098-4
21	276	10.6	529	2	US-08-836-442-3
22	265	10.1	769	1	US-08-243-542-4
23	265	10.1	769	1	US-08-477-407-4
24	265	10.1	769	1	US-08-484-355-4
25	263	10.1	814	4	US-09-813-819-4
26	263	10.1	814	4	US-09-920-048-4
27	263	10.1	855	4	US-09-813-819-2

28	263	10.1	855	4	US-09-920-048-2	Sequence 2, Appli
29	248	9.5	621	4	US-09-026-001A-18	Sequence 18, Appli
30	246	9.4	621	4	US-09-026-001A-6	Sequence 6, Appli
31	239	9.1	613	4	US-09-026-001A-10	Sequence 10, Appli
32	233	8.9	462	4	US-09-411-329C-3	Sequence 3, Appli
33	233	8.9	462	4	US-09-411-329C-17	Sequence 17, Appli
34	231	8.8	464	4	US-09-411-329C-14	Sequence 14, Appli
35	229.5	8.8	470	1	US-08-514-014-2	Sequence 2, Appli
36	229.5	8.8	470	2	US-08-833-823-2	Sequence 2, Appli
37	228.5	8.7	616	4	US-09-608-790-1	Sequence 1, Appli
38	226.5	8.7	470	3	US-08-813-150-2	Sequence 2, Appli
39	224.5	8.6	200	2	US-08-836-442-4	Sequence 4, Appli
40	220.5	8.4	521	4	US-09-026-001A-12	Sequence 12, Appli
41	216	8.3	355	4	US-09-152-060-64	Sequence 64, Appli
42	213.5	8.2	751	2	US-08-836-443-3	Sequence 3, Appli
43	213	8.1	592	4	US-09-026-001A-14	Sequence 14, Appli
44	212	8.1	439	4	US-09-026-001A-8	Sequence 8, Appli
45	209.5	8.0	488	1	US-08-243-542-1	Sequence 1, Appli

## ALIGNMENTS

RESULT 1  
US-09-122-126B-15  
Sequence 15, Application US/09122126B  
Patent No. 6A51525  
GENERAL INFORMATION:  
APPLICANT: Bristol-Myers Squibb Company  
TITLE OF INVENTION: AGGRECAN DEGRADING METALLO PROTEASES  
FILE REFERENCE: DM6909  
CURRENT APPLICATION NUMBER: US/09/122,126B  
CURRENT FILING DATE: 1998-07-24  
NUMBER OF SEQ ID NOS: 21  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 15  
LENGTH: 930  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-122-126B-15

Query Match 97.5%; Score 2549; DB 4; Length 930;  
Best local similarity 99.8%; Pred. No. 3.5e-238;  
Matches 481; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	MLGMASTLLCAFRLLPLAAVGPATPAODKAGOPPTAAAGPPTRRGGEVERAEPGH	60
DB	1	MLGMASTLLCAFRLLPLAAVGPATPAODKAGOPPTAAAGPPTRRGGEVERAEPGH	60
QY	61	PHPLAQRRSKGLVONIDOLYSGGKVGILVYAGGRFTLDERGSGVIGAFVAGGT	120
DB	61	PHPLAQRRSKGLVONIDOLYSGGKVGILVYAGGRFTLDERGSGVIGAFVAGGT	120
QY	121	SAPWRHRSKCFYRGVYDGSFRLAVFDLCCGLDGFPAVKHARYTLKPLIRGFWAEDEKGR	180
DB	121	SAPWRHRSKCFYRGVYDGSFRLAVFDLCCGLDGFPAVKHARYTLKPLIRGFWAEDEKGR	180
QY	181	VYGDGSAIRLLHYTREGFEFLPPRASCETPASPPEAHNAIPASNPGRALASQLLD	240
DB	181	VYGDGSAIRLLHYTREGFEFLPPRASCETPASPPEAHNAIPASNPGRALASQLLD	240
QY	241	QSALSPAGSGGPTQWRRRRRSISRAQVELLLVADASMARLYRGLOHYLLTLASIANR	300
DB	241	QSALSPAGSGGPTQWRRRRRSISRAQVELLLVADASMARLYRGLOHYLLTLASIANR	300
QY	301	LYSHASIEHNIPLAVYKVVVLGDOKSLEVSNAATTLKNFCKWQHONQDDEHYD	360
DB	301	LYSHASIEHNIPLAVYKVVVLGDOKSLEVSNAATTLKNFCKWQHONQDDEHYD	360
QY	361	AILFTFREDLGHHSOPLTGMADVGTICSPERSCAVIEDDGLHAFTVAHEIGHLGLSH	420
DB	361	AILFTFREDLGHHSOPLTGMADVGTICSPERSCAVIEDDGLHAFTVAHEIGHLGLSH	420



NUMBER OF SEQ ID NOS: 31  
 SOFTWARE: Patentin Ver. 2.1  
 SEQ ID NO 9  
 LENGTH: 905  
 TYPE: PRT  
 ORGANISM: Mus musculus ADAMTS-8  
 US-09-369-364A-9

Query Match 28.0%; Score 732.5; DB 4; Length 905;  
 Best Local Similarity 33.5%; Pred. No. 4,4e-62;  
 Matches 177; Conservative 61; Mismatches 165; Indels 105; Gaps 11;

QY 4 GWASLLCAFLPLAAGVPAATPAODKAGPPTAAAMAPRRRQGEVEQERAP---PGH 60  
 DB 9 GMPPLLLLLQLP-----PPPLVCGAPAGCGTGAQASSELVLPRLPS 51  
 QY 61 PHLAQRKRSKGLVONIDOLYSGGKVGILYVAGGRFLDLERGVS-----GLNG 112  
 DB 52 ASELA-----FHLSAFGQGVLRPLAPASFLAPEFKIERLGC 88  
 QY 113 FVPAGGTSAPWRHSHCFYRGTVDSPPSLAVFDLCGIDGFFAVKHARYTLKPLRGP 172  
 DB 89 SGAAGGEG-----LRGFTSGTVNGERESLAAASCAGSGSFLAGEEFTIOPQAGD 144  
 QY 173 -----MAEEKGRVYGDGSAKILHVYTRGFSFELPPRASCEPASTPEAHENA 222  
 DB 145 SLDPQRLQRMGPGGRREDPGLAAEVF-----PLPQGLMEVEMNGGQERS 193  
 QY 223 P-AHNPSPGRALASQLDQSLSPAGSGPQOTWMRRRRRSISRAROVELLVADASMAR 281  
 DB 194 DNEEDKKODEKLKETEDESRKVPPEFGSK-----TRSKRFSEARFVETLLVADASMA 248  
 QY 282 IYGRGLQHTLLASTIANRLYSHASIENHIRLAVYVYVLDGDKDSLEVRKNAATTLKNF 341  
 DB 249 FYGTDLQNHILTVMAARIRYKHPISIRSVNLVYVAVLIVERKNGPEVSDNGGLTRNF 308  
 QY 342 CKWQHONOLDGDEHEHYDAILFTREDLGH-HSCDTLGMADVGTICSPERSCAVIEDD 400  
 DB 309 CSMQRFNKPSPDRHEHYDTAILFTROMFCGECQCDTLGMADVGTICDPKSCSVYNDE 368  
 QY 401 GHAAPYVAHEHGLTGLSHDSCFCEFTFSTEDKRLMSSILTSIDASKPMSKCSAT 460  
 DB 369 GLOAAYTLAHEGHVLSMHPHDSKPCVRLFGPMGRYHMAAPFIHVKTLPMSPCSAVYL 428  
 QY 461 TEFLDDGHNCLLDLP-----KQLLGDYK 486  
 DB 429 TELLDGHDCLLDAPTSVLPPLTGLPGHSTLYELDQCKOLFGRDPR 476

RESULT 5  
 US-09-122-126B-2  
 Sequence 2, Application US/09122126B  
 Patent No. 6451575  
 GENERAL INFORMATION:  
 APPLICANT: Bristol-Myers Squibb Company  
 TITLE OF INVENTION: AGGRECAN DEGRADING METALLO PROTEASES  
 FILE REFERENCE: DM6909  
 CURRENT APPLICATION NUMBER: US/09/122,126B  
 CURRENT FILING DATE: 1998-07-24  
 NUMBER OF SEQ ID NOS: 21  
 SOFTWARE: Patentin version 3.0  
 SEQ ID NO 2  
 LENGTH: 837  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-122-126B-2

Query Match 25.2%; Score 658.5; DB 4; Length 837;  
 Best Local Similarity 33.9%; Pred. No. 5,9e-55;  
 Matches 172; Conservative 63; Mismatches 156; Indels 117; Gaps 15;  
 QY 3 LGMASLLCAFLPLAAGVPAATPAODKAGPPTAAAMAPRRRQGEVEQERAPPGH 62

DB 34 LWMLLLLLALSLP-----SARLASPLPREEIV-----FPE 65  
 QY 63 PLAQRKRSKGLVONIDOLYSGGKVGYL--VYAGGRFLDLERDGSVGIAG----- 113  
 DB 66 KL-----NCSVLPGSGAPARLLCRLQAFETLLLEEDSGVOBELTYOYLQ 114  
 QY 114 VPAGGTSAPWRHSHCFYRGTVDSPPSLAVFDLCG-LDGFFAVKHARYTLKPLRGP 172  
 DB 115 APELLCGAP-----GYTLGTFINGDPESVASLHMDGALLGVLYRGAELHLOPLEG- 168  
 QY 173 WAEKGRVYGDGSAKILHVYTRGFSFELPPRASCEP--ASTPEAH---EHAHNSN 227  
 DB 169 -----TPNSAGSPGAILLRKSPA--- 187  
 QY 228 PSGRALASQLDQSLSPAGSGPQOTWMRRRRRSISRAROVELLVADASMARLYGR 287  
 DB 188 -SGGPMK-----NKAPLGSPPRP--RRKRRASLSRFETLLVADKNAATFGAL 238  
 QY 288 OHYLLTLASTIANRLYSHASIENHIRLAVYVYVLDGDKDSLEVRKNAATTLKPKWQ 347  
 DB 239 KYLLTVMAAARAKRHPISIRNPVSLVYTRVILSGEGEPGVGSAOTLASFAMOKG 298  
 QY 348 HNOIGDDEHYDAILFTREDLGHSCDTLGMADVGTICSPERSCAVIEDDGLHAFT 407  
 DB 299 LNTPEPSDPHEDTALFTFRDLGCVSTCDTLGMADVGTICPARSCAIVEDDGLSFT 358  
 QY 408 VAHEHGLTGLSHDSCFCEFTFSTEDK-LMSILTSIDASKPMSKCSATITREFLDD 466  
 DB 359 AAHELGHYNNLHDSKPCISLNGPLSTRHVAAPMAHAPVDEEWPSCSARFTYFDLN 418  
 QY 467 GAGNCLLDLPKKI-----LGDYKDD 488  
 DB 419 GYGHCLDLPKAPLPLVTPFPKXDYAD 446

RESULT 6  
 US-09-369-364A-11  
 Sequence 11, Application US/09369364A  
 Patent No. 6391610  
 GENERAL INFORMATION:  
 APPLICANT: Apte, Suneel  
 APPLICANT: Hurskainen, Tiina L.  
 TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases  
 FILE REFERENCE: 26473/4007/10-30-00  
 CURRENT APPLICATION NUMBER: US/09/369,364A  
 CURRENT FILING DATE: 1999-08-06  
 NUMBER OF SEQ ID NOS: 31  
 SOFTWARE: Patentin Ver. 2.1  
 SEQ ID NO 11  
 LENGTH: 245  
 TYPE: PRT  
 ORGANISM: Homo sapiens ADAMTS-8  
 US-09-369-364A-11

Query Match 21.7%; Score 567.5; DB 4; Length 245;  
 Best Local Similarity 52.7%; Pred. No. 5,9e-47;  
 Matches 106; Conservative 33; Mismatches 61; Indels 1; Gaps 1;  
 QY 258 RRRRISRAROVELLVADASMARLYGRGLNHLTLASTIANRLYSHASIENHIRLAVY 317  
 DB 17 RTRKRFSEARFVETLLVADASMAAFGADLQNHILTLMSVAARIYKHPISIKNSINIMVYK 76  
 QY 318 VYVLDGDKDSLEVRKNAATTLKPKWQHONHOLDGDEHYDAILFTREDLGHNS-C 376  
 DB 77 VLIVDEKMGPEVSDNGGLTRNFQCMQRRPNQPSDRPHEDTALITRNFQCGEGIC 136  
 QY 377 DTLGMADVGTICSPERSCAVIEDDGLHAFTVAHEHGLTGLSHDSCFCEFTFSTEDK 436  
 DB 137 DTLGVADICTICDPKSCSVIIDEGLQAAHNTLAHEGHVLSMHPHDSKCTRLFGPMCKH 196  
 QY 437 RLMSILTSIDASKPMSKCTS 457

DB 197 HWMAPLEVLHNOTLPMWSPCSA 217

RESULT 7  
US-09-369-364A-13

Sequence 13, Application US/09369364A  
Patent No. 6391610  
GENERAL INFORMATION:  
APPLICANT: Apte, Suneel  
APPLICANT: Hurskainen, Tiina L.  
APPLICANT: Hirshata, Satoshi  
TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases  
FILE REFERENCE: 26473/4007/10-30-00  
CURRENT APPLICATION NUMBER: US/09/369,364A  
CURRENT FILING DATE: 1999-08-06  
NUMBER OF SEQ ID NOS: 31  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 13  
LENGTH: 1882  
TYPE: PRT  
ORGANISM: Homo sapiens ADAMTS-9  
FEATURE:  
NAME/KEY: MOD\_RES  
LOCATION: (468)  
OTHER INFORMATION: Xaa = C  
NAME/KEY: MOD\_RES  
LOCATION: (521)  
OTHER INFORMATION: Xaa = Y  
US-09-369-364A-13

Query Match 21.6%; Score 565.5; DB 4; Length 1882;  
Best Local Similarity 29.6%; Pred. No. 2.1e-45;  
Matches 158; Conservative 66; Mismatches 163; Indels 147; Gaps 16;

QY 3 LGSALLLCAFRLLPLAAGBPATPAGDKAGQPTAAAOAPRRKQEEVO----- 52  
DB 6 VSNATLLTLVLR-DLAEMG-----SPDAANAARKKDLAPRQVKLETTSEYEI 52  
QY 53 -----ERAEPGHPPL-AQRRSKGLQVONIDOLY-----SGCGKGYLYVACGRFL 99  
DB 53 VSPRIYNALGEPPTVHVKRTKRSINSATDPAPAFASSTSSSQAYRLSAFCQGL 112  
QY 100 LDIERGSGVIGFV-----PAGGCT---SAPRNHSHCYRGTVDGSPSLAV 145  
DB 113 FNLTA-----AGFIAPLFTVTLGTPGVNOTKFYSEEAELKHCFKRLCOYOLA--- 164  
QY 146 FDLGCGLDGFPFAVKHARYTLKPLRGPMABEEKGRVYGDGSAKILHYTRREGSFELPP 205  
DB 165 -----HGRH-QPLLR----- 173  
QY 206 RASCETPASTPEAHEHAPHS-----NPSGRAALASQLDOSALSPAGS 250  
DB 174 -----NEHKNHSSKOKKKTARKMGERINLAGDVAALSGLATEAFSAVGNK 220  
QY 251 GPQT-----WRRRRRSISAROVELLVADASHARLYGGLQHYLLTLASIANRLYSHAS 306  
DB 221 TDNTRKRTKRRKRLSYPRFEVLYVADNRVSYHGENLQHYITLMSIVASIKDPS 280  
QY 307 IENHILAVVAVVVGDDKXSLVSKNAATPILKPFCKMOHONHOGDDHEHDAALPT 366  
DB 281 IGLNLIYVNLVLIHNEODGPSISFNAOTTLKNCQOMHSSPGS-----IHMDVALLT 336  
QY 367 REDLQ-GHNSCDTLGMADVGTICSPERSCAVIEDDLHAFTVAHEIGHLGLSHDSKF 425  
DB 337 RODICAHDKDGLGLAELGTICDPYRSCSISDSGLSTAFITAHGLGHVFNPHDNNK 396  
QY 426 CEETGSGTDEKRLMSILTSIDASKPWSKCTSATITFEFLDDGHCNCLDLPK 479  
DB 397 CREE-GVKSPOHVMAPTLPNFYTPMWSKCSRKXYITFEFLDTGCECLNEPESR 449

RESULT 8  
US-09-369-364A-15

Sequence 15, Application US/09369364A  
Patent No. 6391610  
GENERAL INFORMATION:  
APPLICANT: Apte, Suneel  
APPLICANT: Hurskainen, Tiina L.  
APPLICANT: Hirshata, Satoshi  
TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases  
FILE REFERENCE: 26473/4007/10-30-00  
CURRENT APPLICATION NUMBER: US/09/369,364A  
CURRENT FILING DATE: 1999-08-06  
NUMBER OF SEQ ID NOS: 31  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 15  
LENGTH: 874  
TYPE: PRT  
ORGANISM: Mus musculus ADAMTS-9  
US-09-369-364A-15

Query Match 21.4%; Score 560.5; DB 4; Length 874;  
Best Local Similarity 36.9%; Pred. No. 2e-45;  
Matches 133; Conservative 50; Mismatches 126; Indels 51; Gaps 10;

QY 144 AYFDLGGDGFPAVKHARYTLKPLRGPMABEEKGRVYGDGSAKILHYTRREGSFEL 203  
DB 3 AVTSLCSGMGTFRSHDGDFTLEPLQSVDEDEDE-----EONKPHIITRH----- 48  
QY 204 PRASCETPASTPEAHEHAP-----HNSPGRAAL-----ASQLDOSALSPAG--- 248  
DB 49 -----STPQREPTSKHACATSELSKSHSKDKRKIRMKRRRRNSLADVALKSGLAT 102  
QY 249 -----GSGPQT-----W-WRRRRRSISAROVELLVADASHARLYGGLQHYLLTLASI 297  
DB 103 KVLGSGYNOTNTRDWMNHRKTRKFLSYPRFEVVMVADHRMVLHYGANLQHYITLMSI 162  
QY 298 ANRLYSHASLENHILAVVAVVVGDDKXSLVSKNAATPILKPFCKMOHONHOGDDHEE 357  
DB 163 VASTYDSSIGNLINIVNLVLIHNEODGPYINFAOTTLKNCQOMHSSKNTLGG---I 219  
QY 358 HYDAAILFTREDLQ-GHNSCDTLGMADVGTICSPERSCAVIEDDLHAFTVAHEIGHL 416  
DB 220 QHDTAVLVTRREDICRAQDKDGLGLAELGTICDPYRSCSISDSGLSTAFITAHGLGHV 279  
QY 417 GLSHDSKFCETFGSTEDKRLMSILTSIDASKPWSKCTSATITFEFLDDGHCNCLDLP 476  
DB 280 NMPHDSKNCKEE-GVKSPOHVMAPTLPNFYTPMWSKCSRKXYITFEFLDTGCECLNEP 338

RESULT 9  
US-09-130-491-13

Sequence 13, Application US/09130491  
Patent No. 6416974  
GENERAL INFORMATION:  
APPLICANT: Holtzman, Douglas A.  
APPLICANT: Goodearl, Andrew D.J.  
TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83  
FILE REFERENCE: 09404/041001  
CURRENT APPLICATION NUMBER: US/09/130,491  
CURRENT FILING DATE: 1998-08-07  
EARLIER APPLICATION NUMBER: US 60/058,108  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: US 60/054,961  
EARLIER FILING DATE: 1997-08-06  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 13  
LENGTH: 608  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-130-491-13

Query Match 21.0%; Score 549.5; DB 4; Length 608;  
Best Local Similarity 57.7%; Pred. No. 1.3e-44;  
Matches 97; Conservative 27; Mismatches 43; Indels 1; Gaps 1;

```

OY 318 VVLEDDKXKSLSEVSNAAFTLKNPCKMOHNRQJGDDHDEHNDALILFTRDLGCHNSD 377
Db 4 ILIVIDEQGPPEVTSNAALTILFNCNMOKOHNPSPDRADENHDTAILFTRDLGSSQTC 63
OY 378 TLGMADVGTICSPESCACVIEDDGLHAAFTVAHETGHLGLSHDSSKCEEFSGTEDEK 437
Db 64 TLGMADVGTCDPSNCSGVIEDDGLQAAFTTAHETGHLGFENFMHMDAKACASLNGVNDOSH 123
OY 438 LMSLTSLTSDASKPWSKCTSATITFEFLDDHGNCLLDPRKOI-LGCD 484
Db 124 WMASMLSNHDSQPSWCSAYWITSTFLDNGHECLMDPRQNIQPCD 171

RESULT 10
US-09-369-364A-7
; Sequence 7, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Aptec, Sungeel
; APPLICANT: Hurskainen, Taina L.
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 997
; TYPE: PRT
; ORGANISM: Homo sapiens ADAMTS-7
US-09-369-364A-7

Query Match 20.0%; Score 522.5; DB 4; Length 997;
Best Local Similarity 31.3%; Pred. No. 1.2e-41;
Matches 162; Conservative 62; Mismatches 171; Indels 123; Gaps 21;

OY 8 LILCGFRLPLAVGGAAPPAODKAGQ-----PPTAAAAQ-----PRRQGEVQE 53
Db 19 LILCA-----LAFGADGAPAGRATERAALDIYHPRVDGSGFLSYELMPRLRRDVS 74
OY 54 RAEPGHPHPLAQRRRSGVLVONIDOLYSGGKGVLYVYAGGRFLDLERDGSYIGF 113
Db 75 RRDAPAF-YELQGRRELRFNLTAQ-----HLLAPG---EYSETRRRGGIGRAHI 121
OY 114 VPAGGTSAPWRHSH---CYRGTVDGSPR---SLAVFDLCGLDGFPAVKHARTLKP 167
Db 122 -----RAHTPACHILGEVO--DPELEGGIAAISACDGLKGVQLSNEDEYFIEP 167
OY 168 ILRGMAEBEKGKRVYCGDSARILYHYTRGSEFELAPRASCEPASPPEAHENAPAHN 227
Db 168 LDSAP-----ARPGHAQPHVYKKR-----APER----- 191
OY 228 PSGRAALASQLLDOSALDPAG-----GSGPQTW---WRRR-----RSISAROV 269
Db 192 -----LAQRGDSASPSICGVQVYPELESRRERREQQWRPRLRLHRSYSKEKWC 244
OY 270 ELLVADASMAALYLR-GLQHLILLASIANLLYSHASLENIIRLAVKAVVVLGDKDKL 328
Db 245 EFLVADAAKMYEYHGQPOVESYVLTIMNVAGLFFHDPSTGNIHITIVRLVILDEEDL 304
OY 329 EYSKNAATFLTKNFKCMOHOHNOJLGDHNEHYAAILFTREDLCG---HNSCDTLGMADV 386
Db 305 KITHHADNLTLSKFCWKQKSINMKGAHPLRHHDTAILLTRKDLCAAMNRCETLGLSHVAG 364
OY 387 ICSPEKSCAVIEDDGLHAAFTVAHETGHLGLSHDSS-KFCEETFGSTEDKR-LMSSTL 443
Db 365 MCGPRKSCSINEDTGLPLAFTVAHETGHLGSHFGIHDGSGNDCEPV-----GKRPFIMSPQL 419
OY 444 TSIDASKPWSKCTSATITFEFLDDHGNCLLDPRKOI 481
Db 420 LYDAAPLWMSKCSROYITRELDRGWGLCIDDPADDI 457

```

```

RESULT 11
US-09-930-872-4
; Sequence 4, Application US/09930872
; Patent No. 6448388
; GENERAL INFORMATION:
; APPLICANT: Friddele, Carl Johan
; APPLICANT: Hilburn, Erin
; TITLE OF INVENTION: NO. 6448388eal Human Proteases and Polynucleotides Encoding the
; FILE REFERENCE: LEX-0219-USA
CURRENT APPLICATION NUMBER: US/09/930,872
CURRENT FILING DATE: 2001-08-14
PRIOR APPLICATION NUMBER: US 60/225,852
PRIOR FILING DATE: 2000-08-16
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 4
LENGTH: 1224
TYPE: PRT
ORGANISM: homo sapiens
US-09-930-872-4

Query Match      17.7% Score 462.5; DB 4; Length 1224;
Best Local Similarity 29.2%; Pred.No.1.le-35;
Matches 155; Conservative 66; Mismatches 203; Indels 107; Gaps 23;

QY      5 WASLLCAFRFLPLAAGVF-AATPAODKAGOPTAAAAOPRRROG-----EEV 51
         | : : : : : : : : : : : : : : : : : : : : : : : :
DB      15 WMLLAOVAEOAPACAMGPAAAPSPSVPRP-----PRAREGEMKEGYELVASVEV 68

QY      52 QERAPPGHPIHLAORRSKGLYNIDOLYSGGCKVGILYAGGR-FLDLERDGSVGI 110
         | : : : : : : : : : : : : : : : : : : : : : : : :
DB      69 DHRGVYSHEIMHHQRARRAAVAASESL-----HLRLKSGRHDPHYDLRSTSSLVA 120

QY      111 AGFVAAGS-----TSAPWRHSRCFYGTVDGPSRLAVFDLCGDLGFPAVKHAR 162
        | : : : : : : : : : : : : : : : : : : : : : : : :
DB      121 PGFIYOTLGKTGTRKVQLRP---EDPCFYOGSLSHSNMSVALSTCGLSGMTREAD 177

QY      163 YTLKLPLRGPP---WAEEREKRVYGDGSAIRLLHVYTRREGFSFEALPPRASCTEPASTPEA 218
         | : : : : : : : : : : : : : : : : : : : : : : : :
DB      178 YFLRPL---PSHLWS---KLGRA-AQGSSPSHYLYKR---STEPHAPGASVELVTSFTWE 227

QY      219 HHMAAHNSPSGRALLAQO-----LLDO--SALSAPAGSGSGPTW 255
         | : : : : : : : : : : : : : : : : : : : : : : : :
DB      228 LAHQQLHSSDL-RGLPKQHFCGRKKYMDQPKEDLFILPDEYKSL----- 275

QY      256 WRRRRSRISRAQO-----VELLLVADASMARLYG-RGLQHLVLTLASIANRLYSHASIEH 309
        | : : : : : : : : : : : : : : : : : : : : : : : :
DB      276 --RHKRSLLRSHRNDELNAVETLVVVDDKKMKMGNHGHNITTYVTLLNNVSALFKKGITIGC 333

QY      310 HIRLAIVKVVVLGDKDKSLSEVSKNAATTLKNECKMOHOHNOLOGDDHEEHYAAILTFRED 369
         | : : : : : : : : : : : : : : : : : : : : : : : :
DB      334 NINIAIVGILLLEDDOPGLVISIHADHTLISFCOW--OSGLMGKXGTPRH-DHAILLTOLD 390

QY      370 ICG--HNSCDTLGMADVGCISPERSCAVIEDDGAAAFVAHAHEIHLGLSHD-DSKFC 426
         | : : : : : : : : : : : : : : : : : : : : : : : :
DB      391 ICSMNNECDTLGFAPISGMCKSYRSCITINEDTGIGLAFITTAHESGHNFPMGHDOEGCMNC 450

QY      427 EETFGSTEDKRLMSILNIDSARKPWKSCTSATITINEFLDGGHCNLIDLPR 477
         | : : : : : : : : : : : : : : : : : : : : : : : :
DB      451 KRSEGN----IMSPTLAGRNCVGFMSPCSRYLHKFLSTAQAICLDADPK 496

RESULT 12
US-09-930-872-2
; Sequence 2, Application US/09930872
; Patent No. 6448388
; GENERAL INFORMATION:
; APPLICANT: Friddele, Carl Johan
; APPLICANT: Hilburn, Erin
; TITLE OF INVENTION: NO. 6448388eal Human Proteases and Polynucleotides Encoding the
; FILE REFERENCE: LEX-0219-USA
```

```
?
?
?      CURRENT APPLICATION NUMBER: US/09/930,872
?      CURRANT FILING DATE:    2001-08-14
?      PRIOR APPLICATION NUMBER: US 60/225,852
?      PRIOR FILING DATE:    2000-08-16
?      NUMBER OF SEQ ID NOS:   5
?      SOFTWARE: FastSeq for Windows Version 4.0
?      SEQ ID NO 2
?      LENGTH: 491
?      TYPE: PRT
?      ORGANISM: homo sapiens
US-09-930-872-2
```

Query Match	16.98;	Score 441.5;	DB 4;	Length 491;
Best Local Similarity	29.28;	Pred. No. 2.8e-34;		
Matches 151; Conservative	65;	Mismatches 195;	Indels 107;	Gaps 23;

QY	5	WASLLLAFLRPLAAYR--ATPQADKAGOPPTMAAAOQRRRG-----EEV	51
Db	15	WMLLAQVAEQAPACAMKPPAAAPCSBPVPRP-----PRAERGMKEKEIDYSAIEV	68
QY	52	QERAEPPQHPLPLAORRRSGKLYONIDQLYSGGKGYLYVAGGR--FLDLERDSVGI	110
Db	69	DHRDDYSHETIMHQRRRAVAASEVSL-----HLRLKSGRRHPHYDLRTSSILA	120
QY	111	AGFPADGG-----TSAFWRRSHCFRGYTDGSPRLAAYDLGGIDGDFPAAVHAR	162
Db	121	PGFYVOTLQKGTGTSVOTLP--EDRCFYQGSILRSIRNSVALSTCGGLSGMIRTEAD	177
QY	163	YTLKPLLRGP-----WAEERKGVYGDGSAARILHYVTRGGSFEALPPRASCEPPASTPEA	218
Db	178	YFLRPL--PSHLSM--KLGRA-AQSSSPSHVLYKR--STEPHAGASAEVLYVSTWE	227
QY	219	HEHRAHNSPSCRAALASQ-----LLDP--SALSPAGSGSQTW	255
Db	228	LAHQPLHSSDL-RGLGPOKHQFCRCRRKKYMPQPKEDFLIPDPYKCL-----	275
QY	256	WRRRRSISRARQ-----VELLVAADASMARLYG-RGLQHYLLTASIANRLYSHASIEH	309
Db	276	--RHKRSLRLREHNEDELVEVETLVVYDDKKMMQHNHEHNTITVYLLNMVSALEFKDTIGG	333
QY	310	HIRLAAYVVVVYLGDKDKSLEYSKNAATLTLLNFCWQMOHNOHGDHDEHNDALFLTRFD	369
Db	334	NINIAIVGLILBEEOPELVSHHADTLSSFCWM--QSGLMGKDGTRH--DHALLTLGLD	390
QY	370	LCG--HHSCDTLGNADVGTICSPERSCAVIEDDGLAAHFAVNAHRIGHILGLSHD--DSKFC	426
Db	391	ICSMKNPECDTLGAPFISGMCSKYSRCSCTINEDTGLAFTIAHSGHNPFMHIHDEGENMC	450
QY	427	EETFGSTEDKRLMSSILTSDASKPYKCSATITTEFL	464
Db	451	KKSEGN-----IMSPTLAAGRNCVSESPSCSYQYLHKFL	483

RESULT 13  
 US-09-369-364A-5  
 Sequence 5, Application US/09369364A  
 Patent No. 6391610  
 GENERAL INFORMATION:  
 APPLICANT: Apce, Sunee1  
 APPLICANT: Hurskainen, Taina L.  
 APPLICANT: Hirohata, Satoshi  
 TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases  
 FILE REFERENCE: 26473/4007/10-30-00  
 CURRENT APPLICATION NUMBER: US/09/369,364A  
 CURRENT FILING DATE: 1999-08-06  
 NUMBER OF SEQ ID NOS: 31  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 5  
 LENGTH: 859  
 TYPE: PRT  
 ORGANISM: Homo sapiens ADAMTS-6  
 FEATURE:  
 NAME/KEY: MOD\_RES

; LOCATION: (450)  
 ; OTHER INFORMATION: Xaa - L  
 US-09-369-364A-5

Query Match	15.8%;	Score 413;	DB 4;	Length 859;
Best Local Similarity	28.4%;	Pred. No. 3.9e-31;		
Matches 118;	Conservative 69;	Mismatches 131;	Indels 98;	Gaps 16

QY	65	AORRRSKLV----	ONIDOLYGGGKGVGLVAGRRFLLEDPGSGVIGAFVPAAGGTS	122
		:::    :::	:::    :::	
Db	67	SRRRSMDPRIDPOAVSKLF-----	FKLSAYGKHNLNLTLNTDTSVSKHFTYEWGKD	119
QY	122	AP-WRH--RSHCFYRGTV-DGSPRSIAVFDLCGLDGEFAVKAARYLKLPLLRGMAEE	177	
		:::    :::	:::    :::	
Db	120	GPQKKHDFLDONCHTGYLDQDRSTTKVALSNCVGLHGVIATEDEETPLKLN--TTEDS	172	
		:::    :::	:::    :::	
QY	178	KGRVYGGSSAIIIVYTRREGSFALPRPASCETPASTPRENHENAPHSNPSGGAALASQ	237	
		:::    :::	:::    :::	
Db	178	KHFSTYENCHPHYI-----YKKSALDOR-----	HLUDSH-----	206
QY	238	LLDOSALSPAGSGPQTW-----	RRRRSISAROVETLLVADAS	278
		:::    :::	:::    :::	
Db	207	----CGVSDPTRSG-KRPWMLNDPLPLIHOINNTNHHNRKRSVSIEREFETLLVADKM	267	
QY	279	MARLYGR-GLQHYLLTLTASIANRLRYSHASIEHNIRLAVVKKVVLVDGDKSLSEVKNMAT	337	
		:::    :::	:::    :::	
Db	262	MVGHNGKRDIEHYLLSWNIYAKLEROSSLCGNVNIYARLLVLTEDOORNLIEIHNHADS	322	
		:::    :::	:::    :::	
QY	338	LKNFCKWQ-----HQH--NQLGDDHEEYDAAILFTREDLCGHS--	CDTLGMADVSTIC	386
		:::    :::	:::    :::	
Db	322	LDSFCQKWKSTLSHQSDGNTIPRENGIAHNDAVILITRYDICTYKKNKPCGTGLGLASVAGMC	381	
QY	389	SPERSCAVIDEDGLHAATVVAHEGHLGLSHD-----	DSKSC	427
		:::    :::	:::    :::	
Db	382	EPERCSSTNEIDIGSAPTLIEHLVHNENGMHNDIGNSCGRYAKQOONYGSSHNCE	437	

```

1 RESULT 14
2 US-09-491-522-11
3 Sequence 11, Application US/094191522
4 Patent No. 6428998
5
6 GENERAL INFORMATION:
7
8 APPLICANT: Collige, Alain
9 APPLICANT: Lapiere, Charles M.
10 APPLICANT: Prochop, Darwin J.
11 TITLE OF INVENTION: RECOMBINANT N-PROTEINASE
12 TITLE OF INVENTION: AND THE PRODUCTION, METH
13 NUMBER OF SEQUENCES: 17
14
15 CORRESPONDENCE ADDRESS:
16 ADDRESSEE: Pennie & Edmonds, LLP
17 STREET: 1155 Avenue of the Americas
18 CITY: New York
19 STATE: NY
20
21 COUNTRY: USA
22 ZIP: 10036-2811
23
24 COMPUTER READABLE FORM:
25 MEDIUM TYPE: Diskette
26 COMPUTER: IBM Compatible
27 OPERATING SYSTEM: Windows
28 SOFTWARE: FASTSEQ for Windows Version 2.0b
29
30 CURRENT APPLICATION DATA:
31 APPLICATION NUMBER: US/09/491,522
32
33 FILING DATE:
34 CLASSIFICATION:
35 PRIOR APPLICATION DATA:
36 APPLICATION NUMBER: 08/886,333
37
38 FILING DATE:
39
40 ATTORNEY/AGENT INFORMATION:
41 NAME: Abrams, Samuel B
42 REGISTRATION NUMBER: 30,605
43 REFERENCE/DOCKET NUMBER: 8389-0060-999
44
45 TELECOMMUNICATION INFORMATION:
46 TELEPHONE: 650-493-4935

```

TELEFAX: 650-493-5556  
 TELEX: 66141 PENNIE  
 INFORMATION FOR SEQ ID NO: 11:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1205 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-09-491-522-11

Query Match 15.6%; Score 407; DB 4; Length 1205;  
 Best Local Similarity 28.2%; Pred. No. 2.5e-30;  
 Matches 144; Conservative 68; Mismatches 209; Indels 90; Gaps 20;

4 GWASLLC-----AFRLPLAAYGPAATPAODKAGOPPTAAA---AAOPRR--ROGEVE 53  
 6 GAAGLLCPALLLLLLLLPLPADARLAAAAADPPGGPQGGAGARRIAYPRRTAAGRLVSH 65  
 54 RAEPGHPHPLAQRRRRSKGLVONIDOLYSG-----GGKGYLVYAGRRFLDLERDGS 107  
 66 VVSATAPAGVTRRAAPAIQIPGL-----SGSEEDPGGRLFYNTYFGDLHLRLRFNAR 121  
 108 VGIAGFVPRAGGTSAPWRHRS-----HCFRTGYDGSPLSLAV-FDLGGGLDGF 156  
 122 L-----VAPGATVEWQESGATRYEPLGLTYVGVDAESSVALSCDGLAGLI 174  
 157 AVKHARYTLKPLRGPMADK-EKGRYVGGDSARILHVTYREGF---SFEALPPRASCT 211  
 175 RNEEEFFLEPEKGLAAKEADGRVH-----VYHRTFTSPFFLGGQALDGTGSIADS 229  
 212 PASTPEAHENAHANSPSGRAALASOLLDOALSPAGSGSPOTWRRRRRSISRAROVLE 271  
 230 LLSLSRSLGVLDERVNSSR-----RMRHADDDVDNI 264  
 272 LLYADASMARLYG-RGLQHYLLTASIANRLYSHASIEHRIAYVKVYVYG-DKDKSLE 329  
 265 LIGVDSVYQFGTEHYOKYLLTLMNIVMEIYHDESLGHINWVYVRIILLYGKSMSLI 324  
 330 VSKNAATLTKNCKMOHONOLGDDHEEHYDAAILFTREDLGGHSCDTLGMADVGTICS 389  
 325 ETGNPSOSILENCRNAYLQOKPDTDHEDYHDAIFLTRDF---GRSGMGQAPYTGKCH 381  
 390 PERSCAVIEDDGLHAFTVAHEIGHLGLSHD-DSKFC--EETFGSTEDKRLMSILTSI 446  
 382 PVRSCTLNEHDEGFSFAFYVAHEHGLVMEHDOGNRCGDEVRLGS-----IMAPLVQAA 436  
 447 DASKPRSKCTSATITFEFLDDHG-NCILDLP 476  
 437 FHRFMSRCSQOELSRYL--HSYDCLRDLP 464

RESULT 15  
 US-09-369-364A-17  
 ; Sequence 17, Application US/09369364A  
 ; Patent No. 6391610  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Apte, Suneel  
 ; APPLICANT: Hurskainen, Tiina L.  
 ; APPLICANT: Hirahata, Satoshi  
 ; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases  
 ; FILE REFERENCE: 26473/4007/10-30-00  
 ; CURRENT APPLICATION NUMBER: US/09/369,364A  
 ; CURRENT FILING DATE: 1999-08-06  
 ; NUMBER OF SEQ ID NOS: 31  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 17  
 ; LENGTH: 1081  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens ADAMTS-10  
 US-09-369-364A-17

Query Match 14.8%; Score 388; DB 4; Length 1081;  
 Best Local Similarity 26.0%; Pred. No. 1.5e-28;

Matches 136; Conservative 69; Mismatches 192; Indels 126; Gaps 21;  
 25 TPADKAGOPPTAAAAAPRRRQGEVEVERAEP-----GHPHPLAQRRSKGLVQNIQD 79  
 3 TPSSGLKMSSCPWWRAMRSB-----SPRMTTGHCP-----SRHLPLGAP 44  
 80 LYSGGKGYLVYAG--GRFLDLERDGSVGIAGFVAGGCT--SAPWRH--RSNCFYR 133  
 45 RHGHSRVPRLQSGLASTHFLNLTSSRL-LAGRVSVEYVTRGGLAMQRAARPHCLYA 103  
 134 CIVDGSPLSLAV-FDLGGGLDGFFAVKHARTLKLKGP---MAEEKGR--VYGDSA 187  
 104 GHIQOQASSSHVAISTCGLMGLIYADEEYDLEPLHGGPKGSRSPDESGPHCYKRSSL 163  
 188 RLHYVTRGFGFEALPPRASCTPASTPEAHENAHANSPSGRAALASOLLDOALSPA 247  
 164 RHPHLDITACG-----VADKPMKGRP----- 184  
 248 GSGSPOTWWRRR-----RRSISRAROVLELLYADASMARLYG-RG 286  
 185 -----WMLRFLKPPRAPRLGNTERGQPLKRSVSHERYETMDVADKMWVAYHGRD 237  
 287 LQHYLLTASIANRLYSHASIEHRIAYVKVYVYGDKDKSLEVSKNAATLTKNCKMOH 346  
 238 YEQYVLAIMNIVAKLFODSSLSGTVNIIIVTLRLILTEDQPTLEITHAGKSIDSFCMKOK 297  
 347 -----QHNOGDDHEEHYDAAILFTREDLGGHSS--CDTLGMADVGTICSPERSCAVI 397  
 298 STVINISGHGNALPENGVANHDYAVLITRYDICTYNNKPGCTGLGLRMACEVSARELAASM 357  
 398 EDDGLHAFTV--AHIEGHLGLSHD--SKFCEFTFGSTEDKRLMSILTSIDASKPMSK 454  
 358 R--TLAATSVNHCHEIGITFGMNDGYVNSCGAR--GODPAKLAHAHTTKMTNPFVMS 412  
 455 CTSATITFEFLDDHGNCILDLPKQ-----TLGDDYKDDDK 491  
 413 CNRDYITTSFLDSGLGICLNNRPRDFFVPTVAPCAAYDADQ 455

Search completed: April 21, 2003, 14:49:18  
 Job time : 25.4345 secs

**THIS PAGE BLANK (USPTO)**





DR N-PSDB: AAB00726.  
XX New isolated aggrecan degrading metalloproteases - used to develop  
PT products for treating, e.g. osteoarthritis, joint injury, reactive  
PT arthritis, psoriatic arthritis or juvenile rheumatoid arthritis  
XX  
PS Claim 26; Page 68-70; 73pp; English.  
XX  
CC This sequence represents the human aggrecan degrading metalloprotease 1  
CC (ADMP-2). ADMP-1 (AAM75425) and ADMP-2 are novel proteases that cleave  
CC the aggrecan (a major cartilage proteoglycan) between residues  
CC Glu373-Ala374 of the interglobular domain (compared with cleavage between  
CC Asn341-Phe342 by the matrix metalloproteases MMP-1, -2, -3, -7, -8, -9  
CC and -13). ADMP-1 and ADMP-2 were isolated and purified from the  
CC conditioned media of bovine nasal cartilage stimulated by  
CC interleukin-1beta. The purified proteins were partially sequenced and  
CC primers were synthesised based on the resultant amino acid sequences  
CC (AAM00727-X00732 for ADMP-1 and AAM00733-X00736 for ADMP-2). These were  
CC used to isolate the corresponding genes from cDNA. The ADMP polypeptides  
CC can be used for identifying inhibitors of ADMP activity which would  
CC prevent cleavage of the aggrecan core protein, thereby decreasing the  
CC loss of aggrecan from cartilage. Such inhibitors can be used for treating  
CC diseases such as osteoarthritis, joint injury, reactive arthritis, acute  
CC pyrophosphate arthritis (pseudogout), psoriatic arthritis and juvenile  
CC rheumatoid arthritis.  
XX  
SQ Sequence 930 AA:  
  
Query Match 97.5%; Score 2549; DB 20; Length 930;  
Best Local Similarity 99.8%; Pred. No. 4.2e-221;  
Matches 481; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 MLIGNASLILCAFRLLPLAAVGPATPAODKAGPTAAAPARRRGGEEVERAEPGH 60  
DB 1 MLIGNASLILCAFRLLPLAAVGPATPAODKAGPTAAAPARRRGGEEVERAEPGH 60  
QY 61 PHLPLAQRRRSKGLVNDIDLYSGGKGVLYVAGGRFLDLERDGSVGIAGFPVAGGT 120  
DB 61 PHLPLAQRRRSKGLVNDIDLYSGGKGVLYVAGGRFLDLERDGSVGIAGFPVAGGT 120  
QY 121 SAPWHRHSICFRTGVDSGPRSLAVFDLGGDGFVAVYHAYTTLKPLLRGMWAEKGR 180  
DB 121 SAPWHRHSICFRTGVDSGPRSLAVFDLGGDGFVAVYHAYTTLKPLLRGMWAEKGR 180  
QY 181 VYGDSSARILHYTRREGSFELPPRASCETPASTPEAHEHAPAHNSPGRALASQLLD 240  
DB 181 VYGDSSARILHYTRREGSFELPPRASCETPASTPEAHEHAPAHNSPGRALASQLLD 240  
QY 241 OSALSPAGSGSGQOTWRRRRRSISRAROVELLVADASMARLYGKGLHYLLTLASINR 300  
DB 241 OSALSPAGSGSGQOTWRRRRRSISRAROVELLVADASMARLYGKGLHYLLTLASINR 300  
QY 301 LYSNHSIENHILAVVAVVYVLDGDKSLVSKNAATTLKPFCKMHOHQLDDHEHND 360  
DB 301 LYSNHSIENHILAVVAVVYVLDGDKSLVSKNAATTLKPFCKMHOHQLDDHEHND 360  
QY 361 AAILFTREDLCGHSCDTLGMADVGTICSPERSCAVIEDDGLHAFAFVAHETGHLGLSH 420  
DB 361 AAILFTREDLCGHSCDTLGMADVGTICSPERSCAVIEDDGLHAFAFVAHETGHLGLSH 420  
QY 421 DDSKCEETFGSTEDKRLMSSILTSIDASKPWSKCTSATITEFLLDGHGNCLLDPRKQI 480  
DB 421 DDSKCEETFGSTEDKRLMSSILTSIDASKPWSKCTSATITEFLLDGHGNCLLDPRKQI 480  
QY 481 LG 482  
DB 481 LG 482  
  
RESULT 2  
AAB41226  
ID AAB41226 standard; Protein: 929 AA.  
XX

AC AAB41226;  
XX  
DT 08-FEB-2001 (first entry)  
XX  
DE Human ORFX ORF990 polypeptide sequence SEQ ID NO:1980.  
XX  
XX Human; open reading frame; ORFX; detection; cytosolic; hepatotropic;  
XX vulnerable; antiparasitic; antiparkinsonian; neurotropic; neuroprotective;  
XX immunosuppressant; osteopathic; antirheumatic; immunosuppressant; cardiant;  
XX immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
XX hypotensive; dermatological; immunosuppressive; antineoplastic;  
XX antilethal; antibacterial; antifungal; antineuritic; antihypertensive;  
XX antineuritic; gene therapy; cancer; proliferative disorder; hypertension;  
XX neurodegenerative disorder; osteoarthritis; graft vs host disease;  
XX cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
XX cholesterol ester storage; systemic lupus erythematosus; infection;  
XX severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
XX allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
XX bone damage; cartilage damage; antineoplastic disease; coagulation;  
XX thrombosis; contraceptive.  
XX  
XX Homo sapiens.  
XX  
XX WO200058473-A2.  
XX  
XX 05-OCT-2000.  
XX  
XX 31-MAR-2000; 2000WO-US08621.  
XX  
XX 31-MAR-1999; 9905-0127607.  
XX  
XX 02-APR-1999; 9905-0127636.  
XX  
XX 05-APR-1999; 9905-0127728.  
XX  
XX 30-MAR-2000; 2000US-0540763.  
XX  
XX (CUBA-) CURAGEN CORP.  
XX  
XX Shinkets RA, Leach M;  
XX WPI; 2000-602362/57.  
XX  
XX N-PSDB: AAC75435.  
XX  
XX Novel nucleic acids and peptides derived from open reading frame X,  
XX useful for treating e.g. cancers, proliferative disorders,  
XX neurodegenerative disorders and cardiovascular disease -  
XX  
XX Claim 11; Page 1497-1500; 5507pp; English.  
XX  
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX  
XX sequences have activities such as: cytosolic; hepatotropic; vulnerable;  
XX antiparasitic; antiparkinsonian; neurotropic; neuroprotective;  
XX osteopathic; antineuritic; antirheumatic; immunosuppressant;  
XX immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;  
XX antidiabetic; hypotensive; dermatological; immunosuppressive;  
XX antineoplastic; antibacterial; antifungal; antineuritic; antihypertensive;  
XX antineuritic; gene therapy; cancer; proliferative disorder; hypertension;  
XX neurodegenerative disorder; osteoarthritis; graft vs host disease;  
XX cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
XX cholesterol ester storage; systemic lupus erythematosus; infection;  
XX severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
XX allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
XX bone damage; cartilage damage; antineoplastic disease; coagulation;  
XX thrombosis; contraceptive.  
XX  
XX Sequence 929 AA:  
  
Query Match 89.3%; Score 2335.5; DB 21; Length 929;  
Best Local Similarity 91.6%; Pred. No. 8.4e-202;  
XX





Matches 195; Conservative 54; Mismatches 149; Indels 115; Gaps 13;

OY 26 PAODKAGQPT-----AAAAOPRRGGEVQRAEP-----PGHHPPLAQRKRSGL 73  
DB 25 PCSRFGVPPTLLLLAALLAVSALRRSEDEDELVPELEBARHG---TTRRLIAAF 81  
OY 74 VONIDOLYSGGKGVLYVAGRRFLDLERDGSVIGFV-----PAGGTSAPWRRHR 127  
DB 82 DQOLD-----LELRPDSFLAPGFTLQWGRKSGSETPLPETDL 120  
OY 128 SHCFRGTVDSRSLAYFDLGGDGFFAVKNARYTLKPLLRGMAEERKRVYGDGSA 187  
DB 121 AHCFYSGTVNDPSSAALSLCEGVGAFTLLGEAVFTQPL---PAASER-----167  
OY 188 RLIHYTREGFSFEALPPRASCETPASTPEAH-----EHAHANSNPSG 230  
DB 168 -----LATRAPEGKPPAPLQFHLRLRNROGDVGTGCVVDDEP---RPTG 209  
OY 231 RAALASQLLDOSALSPAGSSGPOTWMR-----RRRSISRAROVELL 272  
DB 210 KA-----ETEDDEGTEGDEGPQ--WSPDOPALQGVGPTGTGSIKKRFVSSHRYETM 263  
OY 273 LVADSMARLYRGLOHLLTLTASIANRLYSHASTENHRLAVVAVVYLGDOKSLEYSK 332  
DB 264 LVADGSMAEFHSGSLKHYLLTLFVAARLYKHPISIRNSVSLVYVILVIHDEQKPEVTS 323  
OY 333 NAATTLKFNCKQHOHNOGLGDDEEHYDAAILFTREDLCGHSCDTLGMADVGTICSPER 392  
DB 324 NAALTLLRNFCNNQKOHNPSPDRDAEHYDTAILFTRODLGSGTCTTLGMADVGTICDPSR 383  
OY 333 SCAVIEDGSLAAFTVAHEIGHLLGLSHDSSKFCETGTSFEDKRLMSILTSIDASKPW 452  
DB 384 SCSTVEDGLQAAFTTAHELGHVFNMPHDDAKQACASLNGVNDSHMAMSLNLDHSGPW 443  
OY 453 SKCTSATITTEFLDGHGNCILDLPKROI-IGGD 484  
DB 444 SPCSAYMITSFLDNGHGECLMDKPQNPQLPGD 476

RESULT 6  
AAY49501  
ID AAY49501 standard; Protein: 950 AA.  
XX AAY49501:  
AC  
XX  
DT 10-JAN-2000 (first entry)  
XX  
DE Human METH1 protein.  
XX  
XX Human; METH1; METH2; anti-angiogenic; metalloprotease thrombospondin;  
KW cancer; diagnosis; hyperproliferative disorder; autoimmune disease;  
KW angiogenesis inhibitor; abnormal wound healing; inflammation;  
KW rheumatoid arthritis; psoriasis; endometrial bleeding disorder;  
KW diabetic retinopathy; macula degeneration; haemangioma; detection;  
KW arterial-venous malformation; immune deficiency.  
XX  
OS Homo sapiens.  
XX  
XX WO9937660-A1.  
XX  
XX 29-JUL-1999.  
XX  
XX 22-JAN-1999; 99WO-US01313.  
XX  
XX 23-JAN-1998; 98US-0072298.  
XX  
XX 28-AUG-1998; 98US-0098539.  
XX  
XX (IRUE/) IRUELA-ARISPE L.  
XX  
XX (HAST/) HASTINGS G A.  
XX  
XX (RUBE/) RUBEN S M.  
XX  
XX IrueLA-Arispe L, Hastings GA, Ruben SM.

DR WPI: 1999-590684/50.  
DR N-PSDB; AA232000.  
XX  
PT New isolated metalloprotease thrombospondin polypeptides, useful for  
PT treating hyperproliferative disorders, cancers or autoimmune disorders  
PT  
XX  
XX Claim 10; Fig 1; 457pp; English.  
CC AA232000 and AA232001 encode, and AAY49501 and AAY49502 represent, human  
CC metalloprotease thrombospondin (METH) proteins METH1 and METH2  
CC respectively. METH1 and METH2 have been found to be potent inhibitors of  
CC angiogenesis both in vitro and in vivo. They can be used for treating  
CC cancer and other disorders related to angiogenesis including abnormal  
CC wound healing, inflammation, rheumatoid arthritis, psoriasis,  
CC endometrial bleeding disorders, diabetic retinopathy, some forms of  
CC macula degeneration, haemangiomas, and arterial-venous malformations.  
CC They may be useful in treating deficiencies or disorders of the immune  
CC system, by activating or inhibiting the proliferation, differentiation,  
CC or mobilisation (chemotaxis) of immune cells. The etiology of these  
CC immune deficiencies or disorders may be genetic, somatic, such as  
CC cancer or some autoimmune disorders, acquired (e.g. by chemotherapy or  
CC toxins), or infectious. They can also be used to treat inflammatory  
CC conditions, both chronic and acute conditions. The products can also be  
CC used for detection and diagnosis. AA232002 to AA232080, and AAY49503 to  
CC AAY49511 represent sequences given in the exemplification of the present  
CC invention.  
XX  
XX Sequence 950 AA:  
SO  
Query Match 31.7%; Score 829.5; DB 20; Length 950;  
Best Local Similarity 37.8%; Pred. No. 1.2e-65;  
Matches 194; Conservative 54; Mismatches 150; Indels 115; Gaps 13;

OY 26 PAODKAGQPT-----AAAAOPRRGGEVQRAEP-----PGHHPPLAQRKRSGL 73  
DB 8 PCSRFGVPPTLLLLAALLAVSALGRSEDEDELVPELEBARHG---TTRRLIAAF 64  
OY 74 VONIDOLYSGGKGVLYVAGRRFLDLERDGSVIGFV-----PAGGTSAPWRRHR 127  
DB 65 DQOLD-----LELRPDSFLAPGFTLQWGRKSGSETPLPETDL 103  
OY 128 SHCFRGTVDSRSLAYFDLGGDGFFAVKNARYTLKPLLRGMAEERKRVYGDGSA 187  
DB 104 AHCFYSGTVNDPSSAALSLCEGVGAFTLLGEAVFTQPL---PAASER-----150  
OY 188 RLIHYTREGFSFEALPPRASCETPASTPEAH-----EHAHANSNPSG 230  
DB 151 -----LATRAPEGKPPAPLQFHLRLRNROGDVGTGCVVDDEP---RPTG 192  
OY 231 RAALASQLLDOSALSPAGSSGPOTWMR-----RRRSISRAROVELL 272  
DB 193 KA-----ETEDDEGTEGDEGPQ--WSPDOPALQGVGPTGTGSIKKRFVSSHRYETM 246  
OY 273 LVADSMARLYRGLOHLLTLTASIANRLYSHASTENHRLAVVAVVYLGDOKSLEYSK 332  
DB 247 LVADGSMAEFHSGSLKHYLLTLFVAARLYKHPISIRNSVSLVYVILVIHDEQKPEVTS 306  
OY 333 NAATTLKFNCKQHOHNOGLGDDEEHYDAAILFTREDLCGHSCDTLGMADVGTICSPER 392  
DB 307 NAALTLLRNFCNNQKOHNPSPDRDAEHYDTAILFTRODLGSGTCTTLGMADVGTICDPSR 366  
OY 393 SCAVIEDGSLAAFTVAHEIGHLLGLSHDSSKFCETGTSFEDKRLMSILTSIDASKPW 452  
DB 367 SCSTVEDGLQAAFTTAHELGHVFNMPHDDAKQACASLNGVNDSHMAMSLNLDHSGPW 426  
OY 453 SKCTSATITTEFLDGHGNCILDLPKROI-IGGD 484  
DB 427 SPCSAYMITSFLDNGHGECLMDKPQNPQLPGD 459

RESULT 7  
AAB73549

ID AAB73549 standard; Protein: 950 AA.  
 XX AAB73549;  
 AC  
 XX  
 DT 07-AUG-2001 (first entry)  
 XX  
 DE Human ADAM-type metalloprotease MDT5, SEQ ID NO:4.  
 XX  
 DE Human MDT5; ADAM-type metalloprotease; drug screening;  
 KM A Disintegrin And Metalloprotease; cancer; arthritis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN JP2001017183-A.  
 XX  
 PD 23-JAN-2001.  
 XX  
 PF 09-JUL-1999: 99JP-0196584.  
 XX  
 PR 09-JUL-1999: 99JP-0196584.  
 XX  
 PA (YAMA ) YAMANOUCHI PHARM CO LTD.  
 XX  
 DR WPI: 2001-275950/29.  
 XX  
 DR N-PSDB: AAB70224.  
 XX  
 PT A new metal protease and its preparation for use as an anti-cancer and  
 PT anti-arthritis therapeutic -  
 XX  
 PS Claim 1: Page 12-14; 22pp; Japanese.  
 XX  
 CC The invention relates to the novel human ADAM (A Disintegrin And  
 CC Metalloprotease)-type metalloproteases MDT5 (AAB73549) and MDT5  
 CC (AAB73550). The metalloproteases can be used for the treatment of  
 CC cancer and arthritis. The invention also relates to the genes encoding  
 CC MDT5 and MDT5, vectors and host cells containing the MDT5 or MDT5  
 CC genes, the recombinant production of MDT5 and MDT5, and antibody  
 CC specific for MDT5 or MDT5, and methods of screening for compounds  
 CC which modulate the activity of MDT5 and/or MDT5. The present  
 CC sequence represents human MDT5.  
 CC  
 SQ Sequence 950 AA:  
 Query Match 31.7%; Score 829.5; DB 22; Length 950;  
 Best Local Similarity 37.8%; Pred. No. 1.2e-65;  
 Matches 194; Conservative 54; Mismatches 150; Indels 115; Gaps 13:

DB 307 NAALTLRNFCNMQKHNPSPDRDAEHYDTAILETRDLCGSGOTCDTLGMADVGTGCDPSR 366  
 OY 393 SCVIEDDGLHAAFTVAHEIGHLLGSHDSKPCFETPSTEDKRLMSSILTSIDASKPM 452  
 DB 367 SCVIEDDGLHAAFTVAHEIGHLLGSHDSKPCFETPSTEDKRLMSSILTSIDASKPM 426  
 OY 453 SKCTSAITTEFLDDGHGNCCLDLPKQI-LGCD 484  
 DB 427 SPCSAYMITSFLDNGHGECIMDKPQNPQLPGD 459  
 RESULT 8  
 AAB50002  
 ID AAB50002 standard; Protein: 950 AA.  
 XX  
 AC AAB50002;  
 XX  
 DT 19-MAR-2001 (first entry)  
 XX  
 DE Human METH1.  
 XX  
 KM Human: METH1; metalloprotease; thrombospondin; angiogenesis inhibition;  
 KM cancer therapy; benign tumour; ocular angiogenic disease;  
 KM rheumatoid arthritis; psoriasis; wound healing; endometriosis;  
 KM vasculogenesis; granuloma; hypertrophic scar; nonunion fracture;  
 KM scleroderma, trachoma; vascular adhesion; myocardial angiogenesis;  
 KM coronary collateral; cerebral collateral; arteriovenous malformation;  
 KM ischaemic limb angiogenesis; Osler-Webber syndrome;  
 KM plaque neovascularisation; telangiectasia; haemophilic joint;  
 KM angiodioma; fibromuscular dysplasia; wound granulation;  
 KM Crohn's disease; atherosclerosis; birth control.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200071577-A1.  
 XX  
 PD 30-NOV-2000.  
 XX  
 PF 25-MAY-2000: 2000WO-US14462.  
 XX  
 PR 25-MAY-1999: 99US-0318208.  
 XX  
 PR 20-JUL-1999: 99US-0144882.  
 XX  
 PR 10-AUG-1999: 99US-0147823.  
 XX  
 PR 13-AUG-1999: 99US-0373658.  
 XX  
 PR 22-DEC-1999: 99US-0171503.  
 XX  
 PR 22-FEB-2000: 2000US-0183792.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (SMK-) SMITHKLINE BEECHAM CORP.  
 PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.  
 PA (IRUE-) IRUELA-ARISPE L.  
 PA (HAST/) HASTINGS G A.  
 PA (RUBE/) RUBEN S M.  
 PA (JONAK/) JONAK Z L.  
 PA (TRUL/) TRULLI S H.  
 PA (FORN/) FORNMAID J A.  
 PA (TERR/) TERRETT J A.  
 PI IrueLA-Arispe L, Hastings GA, Ruben SM, Jonak ZL, Trulli SH;  
 PI Fornwald JA, Terrett JA;  
 XX  
 DR WPI: 2001-025136/03.  
 XX  
 DR N-PSDB: AAC90057.  
 XX  
 PT METH1 and METH2 polynucleotides and encoded polypeptides, used to  
 PT inhibit angiogenesis in the treatment of disorders such as cancer,  
 PT rheumatoid arthritis and psoriasis -  
 XX  
 PS Claim 15: Fig 1; 768pp; English.  
 XX  
 CC The present sequence is human METH1 (ME for metalloprotease and TH for  
 CC thrombospondin). METH1 can be used for inhibiting angiogenesis in an  
 CC individual, and for treating cancer, benign tumours, an ocular angiogenic

CC disease, rheumatoid arthritis, psoriasis, delayed wound healing,  
 CC endometriosis, vasculogenesis, granulomas, hypertrophic scars, nonunion  
 CC fractures, scleroderma, trachoma, vascular adhesions, myocardial  
 CC angiogenesis, coronary collaterals, cerebral collaterals, arteriovenous  
 CC malformations, ischemic limb angiogenesis, Osler-Webber syndrome, plaque  
 CC neovascularisation, telangiectasia, haemophilic joints, angiodiroma,  
 CC fibromuscular dysplasia, wound granulation, Crohn's disease or  
 CC atherosclerosis, METH1 can also be used in birth control. METH1 can also  
 CC be used in diagnostic methods for the prognosis of cancer.

XX Sequence 950 AA;

Query Match 31.7%; Score 829.5; DB 22; Length 950;  
 Best Local Similarity 37.8%; Pred. No. 1.2e-65;

Matches 194; Conservative 54; Mismatches 150; Indels 115; Gaps 13;

QY 26 PAQDKAGQPT-----AAAAQRRRQGEVQERAP-----PGHNPPLAQRKRSGL 73  
 DB 8 PGSRSGFVPTLLLAALAVSDALGRSEDEDELVPELERAPGHC---TTRLRLLHAF 64  
 QY 74 VQNDQLYSGGKGYLVYAGGRFLDLERDGSVGIAGFV-----PAGGTSAPMNR 127  
 DB 65 DQDLD-----LELRPDSFLAFGFTLQNVGRKSGSETPLPETDL 103  
 QY 128 SHCFYRGTVDSPRSIAVFDLGGIDGFFAVKHAHYTLPLRGPMAREEGRVYGDGSA 187  
 DB 104 AHCFYSGTVNCDPSSAALSLCEGVGAFYLLGEAYFIQPL---PAASER----- 150  
 QY 188 RILHYTREGEFSEALPPRASCETPASTPEAH-----EHAHNSNPSG 230  
 DB 151 -----LATVAAPGEKPPAPLQFHLRLNRNGDVGTGCVVDDP---RPTG 192  
 QY 221 RAALASQLLDOSALSAPGSGPQTWMR-----RRRRSISRAQVLL 272  
 DB 193 KA-----ETEDDEDETEGDEBPQ--WSPQDPALOGVGOPGTGSTRKRKFSSHHRYVTM 246  
 QY 273 LVADASMARLYGRGLQHYLLTLASIANRLYSHASLENHIRLAVKVVYVLGDKDSLEYK 332  
 DB 247 LVADQSMAEFHGSGLKHYLLTLFVAARLYKHPISRNVSLLVYKILVIHDEQKPEVTS 306  
 QY 333 NAATTLKMFCKMWOHONLGDDEHEHYDAILLFTREDLCGHNSCDTLGMAVGTICSPER 392  
 DB 307 NAALLTRNFCMWOHONPPSDRDAEHYDTAILFTRODLGSGOTCDTLGMAVGTICDPSR 366  
 QY 393 SCAVIEDGLAAFTVAHEIGHLGLSHDDSKFCEETFGSTEDKRLMSITLSDASKPW 452  
 DB 367 SCVYIEDGLQAAFTTAELGHVFMPPHDDAKQOCASLNGVNDSHMASMLNLDHSDPW 426  
 QY 453 SKCTSATTTEFLDDGHCNCLDLPRKQI-LGSD 484  
 DB 427 SPCSAYMITSFLDNGHCECLMDKPNPQLPGD 459

RESULT 9  
 AAM80285  
 ID AAM80285 standard; Protein; 967 AA.  
 XX  
 AC AAM80285;  
 XX  
 DT 19-JAN-1999 (first entry)  
 XX  
 DE Human integrin ligand polypeptide ITGL-TSP.  
 XX  
 KW ITGL-TSP: integrin ligand; angiogenic disease; cancer; atherosclerosis;  
 KW chronic inflammatory disorder; rheumatoid arthritis; tissue remodeling;  
 KW macular degeneration; diabetic retinopathy; Alzheimer's disease; human;  
 KW restenosis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP874050-A2.  
 XX  
 PD 28-OCT-1998.

XX 27-JAN-1998; 98EP-0300575.  
 XX 24-APR-1997; 97US-0845496.

XX (HUMA-) HUMAN GENOME SCI INC.  
 PA (SMIK) SMITHKLINE BEECHAM CORP.  
 PA (SMIK) SMITHKLINE BEECHAM PLC.

PI Fromwald JA, Hastings GA, Jonak ZL, Terrett JA;  
 PI Trull SH;

DR WPI: 1998-544643/47.  
 DR N-PSDB: AAV66508.

PT DNA encoding integrin ligand polypeptide ITGL-TSP - used to treat  
 PT angiogenic diseases, restenosis, Alzheimer's disease and in tissue  
 PT remodeling

PS Claim 11; Pages 6-9; 24pp; English.

CC This represents a human integrin ligand polypeptide ITGL-TSP. ITGL-TSP  
 CC polypeptides can be used in the treatment of angiogenic diseases such as  
 CC cancer, cancer metastasis, chronic inflammatory disorders, rheumatoid  
 CC arthritis, atherosclerosis, macular degeneration or diabetic retinopathy,  
 CC restenosis, Alzheimer's disease and tissue remodeling. They can be used  
 CC to treat a subject in need of enhanced activity or expression of the  
 CC ITGL-TSP polypeptide.

XX Sequence 967 AA;

Query Match 31.7%; Score 829.5; DB 19; Length 967;  
 Best Local Similarity 37.8%; Pred. No. 1.2e-65;

Matches 194; Conservative 54; Mismatches 150; Indels 115; Gaps 13;

QY 26 PAQDKAGQPT-----AAAAQRRRQGEVQERAP-----PGHNPPLAQRKRSGL 73  
 DB 25 PGSRSGFVPTLLLAALAVSDALGRSEDEDELVPELERAPGHC---TTRLRLLHAF 81  
 QY 74 VQNDQLYSGGKGYLVYAGGRFLDLERDGSVGIAGFV-----PAGGTSAPMNR 127  
 DB 82 DQDLD-----LELRPDSFLAFGFTLQNVGRKSGSETPLPETDL 120  
 QY 128 SHCFYRGTVDSPRSIAVFDLGGIDGFFAVKHAHYTLPLRGPMAREEGRVYGDGSA 187  
 DB 121 AHCFYSGTVNCDPSSAALSLCEGVGAFYLLGEAYFIQPL---PAASER----- 167  
 QY 188 RILHYTREGEFSEALPPRASCETPASTPEAH-----EHAHNSNPSG 230  
 DB 168 -----LATVAAPGEKPPAPLQFHLRLNRNGDVGTGCVVDDP---RPTG 209  
 QY 221 RAALASQLLDOSALSAPGSGPQTWMR-----RRRRSISRAQVLL 272  
 DB 210 KA-----ETEDDEDETEGDEBPQ--WSPQDPALOGVGOPGTGSTRKRKFSSHHRYVTM 263  
 QY 273 LVADASMARLYGRGLQHYLLTLASIANRLYSHASLENHIRLAVKVVYVLGDKDSLEYK 332  
 DB 264 LVADQSMAEFHGSGLKHYLLTLFVAARLYKHPISRNVSLLVYKILVIHDEQKPEVTS 323  
 QY 333 NAATTLKMFCKMWOHONLGDDEHEHYDAILLFTREDLCGHNSCDTLGMAVGTICSPER 392  
 DB 324 NAALLTRNFCMWOHONPPSDRDAEHYDTAILFTRODLGSGOTCDTLGMAVGTICDPSR 383  
 QY 393 SCAVIEDGLAAFTVAHEIGHLGLSHDDSKFCEETFGSTEDKRLMSITLSDASKPW 452  
 DB 364 SCVYIEDGLQAAFTTAELGHVFMPPHDDAKQOCASLNGVNDSHMASMLNLDHSDPW 443  
 QY 453 SKCTSATTTEFLDDGHCNCLDLPRKQI-LGSD 484  
 DB 444 SPCSAYMITSFLDNGHCECLMDKPNPQLPGD 476

RESULT 10

ID	AB50011
XX	AA50011 standard; Protein: 968 AA.
AC	
XX	AA50011:
XX	
DT	19-MAR-2001 (first entry)
XX	
DE	Protein; SEQ ID 125.
XX	
KW	Human; METH1: metalloprotease; thrombospondin; angiogenesis inhibition;
KM	cancer therapy; benign tumour; ocular angiogenic disease;
KW	rheumatoid arthritis; psoriasis; wound healing; endometriosis;
KV	vasculogenesis; granuloma; hypertrophic scar; nonunion fracture;
KW	scleroderma, trachoma; vascular adhesion; myocardial angiogenesis;
KX	coronary collateral; cerebral collateral; arteriovenous malformation;
KW	ischemic limb angiogenesis; Osler-Weber syndrome;
KM	plaque neovascularisation; telangiectasia; haemophillic joint;
KW	angiofibroma; fibromuscular dysplasia; wound granulation;
KM	Crohn's disease; atherosclerosis; birth control.
XX	
OS	Homo sapiens.
XX	
PN	WO20071577-A1.
XX	
PD	30-NOV-2000.
XX	
PF	25-MAY-2000; 2000WO-US14462.
XX	
PR	25-MAY-1999; 99US-0318208.
PR	20-JUL-1999; 99US-0144882.
PR	10-AUG-1999; 99US-0147823.
PR	13-AUG-1999; 99US-0373658.
PR	22-DEC-1999; 99US-0171503.
PR	22-FEB-2000; 2000US-0183792.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
PA	(SMIK-) SMITHKLINE BEECHAM CORP.
PA	(BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
PA	(IRUE/) IRUELA-ARISPE L.
PA	(HAST/) HASTINGS G A.
PA	(RUBE/) RUBEN S M.
PA	(JONA/) JONAK Z L.
PA	(TRUL/) TRULLI S H.
PA	(FORN/) FORNWALD J A.
PA	(TERR/) TERRETT J A.
XX	
PI	Iruela-Arispe L, Hastings GA, Ruben SM, Jonak ZL, Trulli SH.
PI	Fornwald JA, Terrett JA;
PT	Forwwald JA, Terrett JA;
PT	Inhibitable angiogenesis in the treatment of disorders such as cancer,
PT	rheumatoid arthritis and psoriasis -
PS	
XX	Claim 15: Pages 759-763; 768pp; English.
XX	
CC	The present invention relates to human METH1 and METH2 (ME for
CC	metalloprotease and TH for thrombospondin; see AB5002 and AB5003).
CC	METH can be used for inhibiting angiogenesis in an individual, and for
CC	treating cancer, benign tumours, an ocular angiogenic disease,
CC	rheumatoid arthritis, psoriasis, delayed wound healing, endometriosis,
CC	vasculogenesis, granulomas, hypertrophic scars, nonunion fractures,
CC	scleroderma, trachoma, vascular adhesions, myocardial angiogenesis,
CC	coronary collaterals, cerebral collaterals, arteriovenous malformations
CC	ischemic limb angiogenesis, Osler-Weber syndrome, plaque
CC	neovascularisation, telangiectasia, haemophillic joints, angiofibroma,
CC	fibromuscular dysplasia, wound granulation, Crohn's disease or
CC	atherosclerosis. METH can also be used in birth control. METH can also
CC	be used in diagnostic methods for the prognosis of cancer. The present
CC	sequence is a protein isolated in the present invention.
XX	
XX	Sequence 968 AA:
XX	

Query Match	31.7%	Score 829.5	DB 22	Length 968
Best Local Similarity	37.8%	Pred. No. 1.2e-65		
Matches 194	Conservative	54	Mismatches 150	Indels 115
			Gaps	13

  

QY	26	PAADKAGQPT-----AAAAQPRRQGEVEQRAEP-----GQHPLPLAQRRSKGL	73
Db	26	PGSRFGVPPTLLLLAAMLAVALSALGRSPEDEELVPELERAGHG--TTRRLNAHF	82
QY	74	VQNDQLYSGGKGYLYYAGGRFLDLERGSVGIAGFV-----PAGGTSAPMRIR	127
Db	83	DOOLD-----LELRDPSFLAGFTLQNVGRKSGSETPLETDL	121
QY	128	SHCFYRGTVDSGPSRLAVFDLGGIDGFFAVYNAHYTLKPLRCGPMAEEREKGRVYGDGSA	187
Db	122	AHCFYSGTVNDGPSAALSLCEYRGCAFYLLGEAVFYQLP--PAASER-----	168
QY	188	RLHYVTRBGSFEFLPPRASCETPASTRPEAH-----EHAPAHNSPG	230
Db	169	-----LATAAPGEKPPAPLQFHLRLRRNRQDVGTCGVVDDEP---RPTG	210
QY	231	RAALASOLLDSALSPAGSGSPQWWR-----RRRSISRAHOVELL	272
Db	211	KA-----ETEDDEGETEGDEDEGPO--WSPQDPAALQGVGQPTGSTRKRKFVSHRYVETM	264
QY	273	LVADASMRILYGRGLQHYLLTLASTANRLYSHASTENHTRLAVVVVVYLGDKDKSLEYSK	332
Db	265	LVADQSMAFHFGSGKLHYLLTFVVAARLYKRPISRLNSVLVVYVILYIHDKQKPEVYS	324
QY	333	NAATLTKNFCKWQHONQDGDDEHYDAAILFTREBLCGHHSCOTLGMADVGTCSPEPR	392
Db	335	NAALTLRNFCKWQKOHNPSPDRDAEHYDALFTFRQDLGSGQTCPTLGMADVGTCDPSR	384
QY	393	SCAVIEDDGLAAFTVAHFIHGLLGLSHDDSKFCETPGSTEDKRLMSILTSIDASKPM	452
Db	385	SCSVIEDDGLQAAFTTAHELGHVFPMHPDDAKQCASLNGVNDSSHMAWSMLNLDHSDPM	444
QY	453	SKCTSATTTTEFLDDGHGNCILDLPRKQI-IGCD	484
Db	445	SFCSAYMITSFLLDNCHGECIMDKPQNPQLPQPD	477

  

RESULT 11
AAB21265
ID AAB21265 standard; Protein: 896 AA.
AC AAB21265;
XX
DT 23-FEB-2001 (first entry)
XX
DE Mouse metalloproteinase ADAMTS-1.
XX
KW Mouse; ADAMTS-1; metalloproteinase; ADAM;
KM a disintegrin and metalloproteinase domain; thrombospondin domain;
KM vaccine; nootropic; neuroprotective; antiParkinsonian;
KM cerebroprotective; cytoprotective; antiarthritic; immunosuppressive;
KM Alzheimer's disease; Parkinson's disease; stroke; cancer; arthritis;
KM autoimmune disease; brain tumour; brain injury.
XX
OS Mus musculus.
XX
PN WO200053774-A2.
XX
PD 14-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US06237.
XX
PR 08-MAR-1999; 99US-0264585.
XX
PA (NEUR-) NEUROCRINE BIOSCIENCES INC.
XX
PI Keiner GS, Clark M, Maki RA;
XX





```

XX  Claim 11; Page 297-300; 380pp; English.
PS
XX
CC  This sequence represents a secreted human protein encoded by the gene
CC  clone detailed in the descriptor line. The gene can be used to generate
CC  fusion proteins by linking to the gene to a human immunoglobulin Fc
CC  protein (e.g. AAX04302) for increasing the stability of the fused
CC  protein as compared to the human protein only.
CC  The invention relates to 86 novel genes and their fragments (nucleic
CC  acid sequences: AAX04311-X04410; amino acid sequences AAW78126-W78225)
CC  which are useful for preventing, treating or ameliorating medical
CC  conditions e.g. by protein or gene therapy. Also, pathological
CC  conditions can be diagnosed by determining the amount of the new
CC  polypeptides in a sample or by determining the presence of mutations in
CC  the new polynucleotides. Specific uses are described for each of the 86
CC  polynucleotides, based on which tissues they are most highly expressed in
CC  (see AAX04311 for described uses).
XX
XX  Sequence 967 AA:
SQ
Query Match 31.1%; Score 814; DB 20; Length 967;
Best Local Similarity 38.5%; Pred. No. 3e-64;
Matches 187; Conservative 53; Mismatches 138; Indels 108; Gaps 12:
OY 46 ROGEEVQERAE-----PGHPHPLAQRNRKGLVQNDQLYSGGKGVLYVYAGRRFL 100
DB 52 RPESEDEELVPELBERAGHG---TTRLRLHAFDQOLD-----L 87
OY 101 DLEBGSGVINGV-----PAGGTSAPWRHSHCFRGTVDSGPRSLAVFDLGGCIGD 154
DB 88 ELRPPSSFLAPFTLQWNGRKSSETPLEPTDLAHCFYSGVGNPSSAAALSLCEGVKG 147
OY 155 FFAVNHARTLPLRLGPMAREEKGKRVYGDGSAKILHYTRGFSFEALPPRASCETPAS 214
DB 148 AVELLGEAVFIQPL---PAASER-----LXTAPEGKPPA 179
OY 215 TPEAH-----EHAPHSNPSGRALASQLDLSALSPAGSGGQTMWR 257
DB 180 PLQFHLRLNRNGDVGTCGVVDDEP---RPTGKA---ETEDDEGTEGDEGPO--WS 230
OY 258 -----RRRSISRAROVELLVAADASMAFLYGRGLOHYLLTLASIAN 299
DB 221 PQDPALQGVGPTGTGSRKRRFVSHRYETMLVADDSMAFHSGLKHYLLTFESVAA 290
OY 300 RLYSHASTENHRIALVAVKVVVLGDKDKSLVSKNAATTLKNCCKWQHONOLGDDHEHY 359
DB 291 RLKXHPXIRNSVSLVYVILVHDEQKPEVTSNAALTLRNCNQKQHNPPSDRAEHY 350
OY 360 DAALFTREDLCGHSCTLCGMADVCTICSPERSCAVIEDGDLHAFTVAHEIGHLLGLS 419
DB 351 DTAILEFTMODLCGSGTCTDCLGMADVCTIDPSRCSVIEDDGLQAFTTAHGLGHVHMP 410
OY 420 HDDSKFCEETFGSTEDKRLMSILTSIDASKRWSKCTSATITEFLDDHGNCLDLPRKQ 479
DB 411 HDDAQCASLKNVNDSHMASLSNLDHSQWSPCSAYMTTSFLDNGHGECLMDKPNP 470
OY 480 I-LGCD 484
DB 471 IQLPGD 476

```

```

KW organ shaping; sterility; cancer metastasis.
XX
XX Mus sp.
XX WO9961656-A2.
XX
XX 02-DEC-1999.
XX
XX 28-MAY-1999; 99WO-US11918.
XX
XX 29-MAY-1998; 98US-0087170.
XX 13-APR-1999; 99US-0129023.
XX
XX (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
XX Kimble JE, Bellocch RH;
XX
XX WPI; 2000-072633/06.
XX
XX Identifying modulators of proteins containing metalloprotease and
XX metalloprotease domains, potentially useful for controlling cell
XX migration and organ shaping
XX
XX Disclosure: Flg 1C; 60pp; English.
XX
XX The present sequence represents a murine ADAMTS-1 protein. ADAMTS-1 is
XX a metalloprotease. The specification describes another related
XX metalloprotease, a C. elegans GON-1 protein, that lacks a transmembrane
XX domain and possesses a predicted metalloprotease domain between residues
XX 269-436. In C. elegans hermaphrodites, GON-1 is required for migration of
XX two distal tip cells to produce elongated tubes, whereas in males, GON-1
XX is required for migration of a single linker cell to produce a single
XX elongated tube. The protein is used in the method of the invention. The
XX specification describes a method for identifying a modulator of a
XX protein that contains a metalloprotease domain and a thrombospondin
XX domain. The method comprises treating a target organism, having a
XX developing gonadal cell that is responsive to the protein, with a test
XX compound, and determining any change in migration or shape of the cell
XX attributable to the test compound. The compounds identified are
XX potential therapeutic modulators of abnormal cell migration and organ
XX shaping, e.g. for rendering animals (specifically nematodes) sterile
XX and for inhibiting cancer metastases.
XX
XX Sequence 950 AA:
SQ
Query Match 30.8%; Score 804.5; DB 21; Length 950;
Best Local Similarity 36.9%; Pred. No. 2.1e-63;
Matches 190; Conservative 61; Mismatches 151; Indels 113; Gaps 13:
OY 1 MLGWASILLCAFRILPLAVGPAATPAQDKAGOPTAAAAAOPRRRGEEV---QERAE 56
DB 18 ML-----LLASITMLLCARCAHGRPTE-----EDDELVLPSLERA- 54
OY 57 PGGHHPPLAQRNRKGLVQNDQLYSGGKGVLYVYAGRRFLDLEDDSGVINGF--- 113
DB 55 -PGHDSSTTRLR-----IDAEGQULHLKLDQSGFLAPGFLQ 91
OY 114 -VPAGGTSA-----PMRHRSHCFRGTVDSGPRSLAVFDLGGDGFPAVNHARTLPL 168
DB 92 TVGRSPGSEAGHLDPTGLALHCFYSGVNGDPSAAALSLCEGVGARYLLOGEEFTIP- 150
OY 169 LRGPMAREEKGKRVYGDGSAKILHYTRGFSFEALPPRASCETPASTPEAHHAHNSP 228
DB 151 -----AGVATERLPAVPERESSARQFH---TLRRRR 181
OY 229 SGRALASQLDLSALSPAGSGPOT-----WWR-----RRRSISR 265
DB 182 RSGSGAKGVMDDETL-PTSDSRPESQNTRNQWVYRDPDPODAGKPSGSGSRKRRFVS 240
OY 266 ARQVELLVAADASMAFLYGRGLOHYLLTLASIANRLYSHASTENHRIALVAVKVVVLGDKD 325
DB 241 PRYVETMLVADDSMAFHSGLKHYLLTFESVAAAFYHHPISRLSVYVVKILVITYEQ 300

```



XX The invention relates to an isolated, enriched, or purified protease  
CC polypeptide (I) and polynucleotide (II) encoding (I). (I) may be used to  
CC screen for substances (S) that may modulate its activity. Administering  
CC S (which modulates protease activity in vitro) may be used to treat a  
CC disease or disorder selected from cancers (e.g., of tissues, of blood or  
CC hematopoietic origin, of the breast, colon, lung, prostate, cervical,  
CC brain, ovarian, bladder or kidney), immune-related diseases and  
CC disorders, cardiovascular disease, brain or neuronal-associated diseases  
CC (e.g., central or peripheral nervous system diseases, migraine, pain,  
CC sexual dysfunction, mood disorders, attention disorders, cognition  
CC disorders, hypertension, hyperkinesia, psychotic disorders, neurological  
CC disorders and dyskinesias), metabolic disorders and inflammatory  
CC disorders. (I) may also be useful as a diagnostic tool for a disease or  
CC disorder such as those above. AAU72876-AAU72910 represent human  
CC protease amino acid sequences of the invention.

SQ Sequence 928 AA:

Query Match 29.5%; Score 770.5; DB 23; Length 928;

Best Local Similarity 42.8%; Pred. No. 2,4e-60;

Matches 180; Conservative 49; Mismatches 133; Indels 59; Gaps 12;

OY 82 SGGGKGYLYVAGRRFLDLLENGSVGIAGF-----VPAG--GTSAPWRHSHCFTR 133  
DB 73 SCDQGLFQITAFQEDFEYLHLPDAQFLAPAFSTEHLGVPLQGLTGSSDLR---RCFYS 129  
OY 134 GTVDGSPRLAVFDLGGDLGDFPAVKHARTLTPLRGPAEKEGKRVYGDGSARILHYV 193  
DB 130 GDVNAEPDSFAVSLGGLGAGFYGAELVISPL---PMASAPAAQNSQGA----HLL 182  
OY 194 TREGSFELPPRASCETPASTEPAHEHAPAHNSPGRALAS---QLDQGLSP--- 246  
DB 183 QRRG-----VPGGPS-----GDPTRCGVAGSNPAIL--RALDPYKP 218  
OY 247 -AGSGPQTMRRRRS-----ISRROYELLVADASMARLYXGRLQHYLLTLASIAN 299  
DB 219 RRAGFES--RSRRRSGRARRFVSIPRYVETLVADESWKFPGADLEHYLLTLATAA 275  
OY 300 RLYSHASIEHILAVLVKVVLDKDKSLVSKNAATTLKPFCKMOHONQDGDHEEHY 359  
DB 276 RLYRHPSILNPIVIYVVKVLLBDRSGPRVTGNALTLRNFCAKOKKLKNSDKHPEY 335  
OY 360 DAAILFTREDLCGHSCDTLGMADVCTICSPERSCAVIEDDGLHAAPVAHEIGHLGLS 419  
DB 336 DVAILEFTRODLCATTCDTLGMADVCTMCDPKRSCSVIEDDGLPSAFTAHEIGHVENMP 395  
OY 420 HDSSKFCETFGSTEDKRLMSSILTSIDASKPKSKTSATITFEFLDGHGNCLLDLPKQ 479  
DB 396 HDNVKVCCEVFGLRANHHMSPFTLIQIDRANPWSACSAALITDPLDSGHGDCLLDQPSKP 455  
OY 480 I 480  
DB 456 I 456

Search completed: April 21, 2003, 14:44:54  
Job time : 42.6294 secs

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 14:30:17 ; Search time 15.2332 Seconds

(without alignments)  
1217.071 Million cell updates/sec

Title: US-10-050-200-8

Perfect score: 2370

Sequence: 1 MSOTGSHPGRLAGRLMCA.....PEAPLPLVPTGDKDDDDKG 447

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt-40:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2313	97.6	837	1 AT54_HUMAN	Q75173 homo sapien
2	1145	48.3	630	1 AT54_RAT	Q9ESP7 rattus norv
3	979	41.3	968	1 AT51_MOUSE	P97857 mus musculu
4	977.5	41.2	967	1 AT51_RAT	Q9WUQ1 rattus norv
5	946	39.9	967	1 AT51_HUMAN	Q9UNH8 homo sapien
6	863.5	36.4	890	1 AT58_HUMAN	Q9UNP7 homo sapien
7	846.5	35.7	905	1 AT58_MOUSE	P57110 mus musculu
8	661	27.9	930	1 AT53_MOUSE	Q9U001 mus musculu
9	651.5	27.4	930	1 AT53_HUMAN	Q9UNA0 homo sapien
10	649.5	26.7	1629	1 AT59_HUMAN	Q9P2N4 homo sapien
11	633	26.7	245	1 AT54_BOVIN	Q9LT93 bos taurus
12	534.5	22.6	997	1 AT57_HUMAN	Q9UKP4 homo sapien
13	493.5	20.8	1593	1 AT12_HUMAN	P58397 homo sapien
14	470.5	19.9	1077	1 AT10_HUMAN	Q9H324 homo sapien
15	401	16.9	860	1 AT56_HUMAN	Q9UKP5 homo sapien
16	364.5	15.4	1211	1 AT52_HUMAN	Q95450 h adamts-2
17	360.5	15.2	1305	1 AT52_BOVIN	P79331 b adamts-2
18	329.5	13.9	1205	1 AT53_HUMAN	Q15072 homo sapien
19	249.5	10.5	903	1 AD12_MOUSE	Q61824 mus musculu
20	246	10.4	824	1 AD08_HUMAN	P78325 homo sapien
21	236.5	10.0	480	1 DISA_TRIGA	P15503 trimeresuru
22	225.5	9.5	418	1 HRTE_CROAT	P34182 crocalus at
23	224.5	9.5	414	1 HRTE_CROAT	P15167 crocalus at
24	218.5	9.2	909	1 AD12_HUMAN	Q94384 homo sapien
25	217	9.2	813	1 AD33_HUMAN	Q9B211 homo sapien
26	213.5	9.0	920	1 AD19_MOUSE	Q35674 mus musculu
27	212.5	9.0	478	1 DISR_AGRAR	P30403 agkistrodon
28	208	8.8	413	1 ACIA_AGRAR	Q9PW05 agkistrodon
29	203	8.6	797	1 AD33_MOUSE	Q92349 mus musculu
30	202	8.5	571	1 DIS1_BOVIA	P30431 bothrops ja
31	202	8.5	776	1 AD28_MACRA	Q9X516 macera fasc
32	196.5	8.3	773	1 AD11_MOUSE	Q9TIV4 mus musculu
33	189.5	8.0	207	1 AT55_BOVIN	Q9LT92 bos taurus

34	189	8.0	826	1 AD08_MOUSE	Q05910 mus musculu
35	189	8.0	956	1 AD19_HUMAN	Q9H013 homo sapien
36	188.5	8.0	774	1 AD28_MOUSE	Q9J106 mus musculu
37	187	7.9	814	1 AD15_HUMAN	Q13444 homo sapien
38	182.5	7.7	769	1 AD11_HUMAN	Q75078 homo sapien
39	178	7.5	775	1 AD28_HUMAN	Q9UKQ2 homo sapien
40	173.5	7.3	816	1 AD15_RAT	Q9QV00 r adam 15 p
41	167	7.0	200	1 HR12_LACMU	P22796 lachesis mu
42	164	6.9	761	1 AD24_MOUSE	Q9R160 mus musculu
43	163	6.9	203	1 ADAM_CROAD	P34179 crocalus ad
44	162	6.8	203	1 FIBR_AGRAR	P28891 agkistrodon
45	159	6.7	729	1 AD21_MOUSE	Q9J176 mus musculu

## ALIGNMENTS

RESULT 1  
ID AT54\_HUMAN STANDARD: PRT; 837 AA.  
AC Q75173; Q9UNH8; 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE ADAMTS-4 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 4) (ADAM-TS 4) (ADAM-TS4) (Aggrecanase 1)  
DE (ADMP-1).  
GN ADAMTS4 OR KIA0688.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=98403880; PubMed=9734811;  
RA Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;  
RT Prediction of the coding sequences of unidentified human genes. X. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro.;  
RL DNA Res. 5:169-176(1998).  
[2]  
RX SEQUENCE FROM N.A.  
RX MEDLINE=99286303; PubMed=10356395;  
RA Tortorella M.D., Burd T.C., Pratta M.A., Abbaszade I., Hollis J.M., Liu R.-Q., Rosenfeld S.A., Copeland R.A., Decicco C.P., Wynn R., Rockwell A., Yang F., Duke J.L., Solomon K., George H., Bruckner R., Nagase H., Itoh Y., Ellis D.M., Ross H., Wiswall B.H., Murphy K., Hillman M.C., Jr., Hollis G.F., Newton R.C., Magolda R.L., Trzaskos J.M., Arner E.C.;  
RT Purification and cloning of aggrecanase-1: a member of the ADAMTS family of proteases.;  
RL Science 284:1664-1666(1999).  
[3]  
RX SEQUENCE FROM N.A.  
RX Sawaji Y., Nagase H., Saklatvala J., Clark A.R.;  
RT ADAMTS-4 genomic locus.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
[4]  
RX PARTIAL SEQUENCE, AND CHARACTERIZATION.  
RX MEDLINE=20400518; PubMed=10827174;  
RA Tortorella M.D., Pratta M.A., Liu R.-Q., Abbaszade I., Ross H., Burd T.C., Arner E.C.;  
RT The thrombospondin motif of aggrecanase-1 (ADAMTS-4) is critical for aggrecan substrate recognition and cleavage.;  
J. Biol. Chem. 275:25791-25797(2000).  
CC FUNCTION: CLEAVES AGGREGAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE INVOLVED IN ITS TURNOVER. MAY PLAY AN IMPORTANT ROLE IN THE DESTRUCTION OF AGGREGAN IN ARTHRITIC DISEASES. COULD ALSO BE A ONE CRITICAL FACTOR IN THE EXACERBATION OF NEURODEGENERATION IN ALZHEIMER'S DISEASE.  
CC CATALYTIC ACTIVITY: Cleaves aggrecan at the 392-Glu-1-Ala-393

QY	121	GAPEPTYLGTINQPEESASLHMGGALLGYLOXRGAEHLHQPLEGPPNSAGPGAH	180
QY	121	GAPEPTYLGTINQPEESASLHMGGALLGYLOXRGAEHLHQPLEGPPNSAGPGAH <td>180</td>	180
Db	121	CAPEPTYLGTINQPEESASLHMGGALLGYLOXRGAEHLHQPLEGPPNSAGPGAH	180
QY	181	LRRKSPASGQGPCMVKVAPLPGSPSPDRPRRAKFAFSLSRFETLVVADDKMAFHCGLKR	240
Db	181	LRRKSPASGQGPCMVKVAPLPGSPSPDRPRRAKFAFSLSRFETLVVADDKMAFHCGLKR	240
QY	241	YLTLVMAAAAFKRPISLRNPVSLVVTFLVTLGSGEEPQVQPSAAQTLRSACANORGLN	300
Db	241	YLTLVMAAAAFKRPISLRNPVSLVVTFLVTLGSGEEPQVQPSAAQTLRSACANORGLN	300
QY	301	TPEDSDPHFDPTALILEFTQDLCGVSTCCBTLGMADVGTVCDDPARSCAIVEDDGLQSAFTAA	360
Db	301	TPEDSDPHFDPTALILEFTQDLCGVSTCCBTLGMADVGTVCDDPARSCAIVEDDGLQSAFTAA	360
QY	361	HELGHVFMKLDNKRPCISLNGPLSTSRHWAPVMAHVDPPEPSPSCARFTFDLDNGY	420
Db	361	HELGHVFMKLDNKRPCISLNGPLSTSRHWAPVMAHVDPPEPSPSCARFTFDLDNGY	420
QY	421	GHCLLDKPEAPLHLPTVTGDKDD	444
Db	421	GHCLLDKPEAPLHLPTVTGDKDD	444

RESULT 2

QY	121	GAPEPTYLGTINQPEESASLHMGGALLGYLOXRGAEHLHQPLEGPPNSAGPGAH	180
QY	121 <td>GAPEPTYLGTINQPEESASLHMGGALLGYLOXRGAEHLHQPLEGPPNSAGPGAH</td> <td>180</td>	GAPEPTYLGTINQPEESASLHMGGALLGYLOXRGAEHLHQPLEGPPNSAGPGAH	180
Db	121 <td>CAPEPTYLGTINQPEESASLHMGGALLGYLOXRGAEHLHQPLEGPPNSAGPGAH</td> <td>180</td>	CAPEPTYLGTINQPEESASLHMGGALLGYLOXRGAEHLHQPLEGPPNSAGPGAH	180
QY	181 <td>LRRKSPASGQGPCMVKVAPLPGSPSPDRPRRAKFAFSLSRFETLVVADDKMAFHCGLKR</td> <td>240</td>	LRRKSPASGQGPCMVKVAPLPGSPSPDRPRRAKFAFSLSRFETLVVADDKMAFHCGLKR	240
Db	181 <td>LRRKSPASGQGPCMVKVAPLPGSPSPDRPRRAKFAFSLSRFETLVVADDKMAFHCGLKR</td> <td>240</td>	LRRKSPASGQGPCMVKVAPLPGSPSPDRPRRAKFAFSLSRFETLVVADDKMAFHCGLKR	240
QY	241 <td>YLTLVMAAAAFKRPISLRNPVSLVVTFLVTLGSGEEPQVQPSAAQTLRSACANORGLN</td> <td>300</td>	YLTLVMAAAAFKRPISLRNPVSLVVTFLVTLGSGEEPQVQPSAAQTLRSACANORGLN	300
Db	241 <td>YLTLVMAAAAFKRPISLRNPVSLVVTFLVTLGSGEEPQVQPSAAQTLRSACANORGLN</td> <td>300</td>	YLTLVMAAAAFKRPISLRNPVSLVVTFLVTLGSGEEPQVQPSAAQTLRSACANORGLN	300
QY	301 <td>TPEDSDPHFDPTALILEFTQDLCGVSTCCBTLGMADVGTVCDDPARSCAIVEDDGLQSAFTAA</td> <td>360</td>	TPEDSDPHFDPTALILEFTQDLCGVSTCCBTLGMADVGTVCDDPARSCAIVEDDGLQSAFTAA	360
Db	301 <td>TPEDSDPHFDPTALILEFTQDLCGVSTCCBTLGMADVGTVCDDPARSCAIVEDDGLQSAFTAA</td> <td>360</td>	TPEDSDPHFDPTALILEFTQDLCGVSTCCBTLGMADVGTVCDDPARSCAIVEDDGLQSAFTAA	360
QY	361 <td>HELGHVFMKLDNKRPCISLNGPLSTSRHWAPVMAHVDPPEPSPSCARFTFDLDNGY</td> <td>420</td>	HELGHVFMKLDNKRPCISLNGPLSTSRHWAPVMAHVDPPEPSPSCARFTFDLDNGY	420
Db	361 <td>HELGHVFMKLDNKRPCISLNGPLSTSRHWAPVMAHVDPPEPSPSCARFTFDLDNGY</td> <td>420</td>	HELGHVFMKLDNKRPCISLNGPLSTSRHWAPVMAHVDPPEPSPSCARFTFDLDNGY	420
QY	421 <td>GHCLLDKPEAPLHLPTVTGDKDD</td> <td>444</td>	GHCLLDKPEAPLHLPTVTGDKDD	444
Db	421 <td>GHCLLDKPEAPLHLPTVTGDKDD</td> <td>444</td>	GHCLLDKPEAPLHLPTVTGDKDD	444

RESULT 2

QY	121	GAPEPTYLGTINQPEESASLHMGGALLGYLOXRGAEHLHQPLEGPPNSAGPGAH	180
QY	121 <td>GAPEPTYLGTINQPEESASLHMGGALLGYLOXRGAEHLHQPLEGPPNSAGPGAH</td> <td>180</td>	GAPEPTYLGTINQPEESASLHMGGALLGYLOXRGAEHLHQPLEGPPNSAGPGAH	180
Db	121 <td>CAPEPTYLGTINQPEESASLHMGGALLGYLOXRGAEHLHQPLEGPPNSAGPGAH</td> <td>180</td>	CAPEPTYLGTINQPEESASLHMGGALLGYLOXRGAEHLHQPLEGPPNSAGPGAH	180
QY	181 <td>LRRKSPASGQGPCMVKVAPLPGSPSPDRPRRAKFAFSLSRFETLVVADDKMAFHCGLKR</td> <td>240</td>	LRRKSPASGQGPCMVKVAPLPGSPSPDRPRRAKFAFSLSRFETLVVADDKMAFHCGLKR	240
Db	181 <td>LRRKSPASGQGPCMVKVAPLPGSPSPDRPRRAKFAFSLSRFETLVVADDKMAFHCGLKR</td> <td>240</td>	LRRKSPASGQGPCMVKVAPLPGSPSPDRPRRAKFAFSLSRFETLVVADDKMAFHCGLKR	240
QY	241 <td>YLTLVMAAAAFKRPISLRNPVSLVVTFLVTLGSGEEPQVQPSAAQTLRSACANORGLN</td> <td>300</td>	YLTLVMAAAAFKRPISLRNPVSLVVTFLVTLGSGEEPQVQPSAAQTLRSACANORGLN	300
Db	241 <td>YLTLVMAAAAFKRPISLRNPVSLVVTFLVTLGSGEEPQVQPSAAQTLRSACANORGLN</td> <td>300</td>	YLTLVMAAAAFKRPISLRNPVSLVVTFLVTLGSGEEPQVQPSAAQTLRSACANORGLN	300
QY	301 <td>TPEDSDPHFDPTALILEFTQDLCGVSTCCBTLGMADVGTVCDDPARSCAIVEDDGLQSAFTAA</td> <td>360</td>	TPEDSDPHFDPTALILEFTQDLCGVSTCCBTLGMADVGTVCDDPARSCAIVEDDGLQSAFTAA	360
Db	301 <td>TPEDSDPHFDPTALILEFTQDLCGVSTCCBTLGMADVGTVCDDPARSCAIVEDDGLQSAFTAA</td> <td>360</td>	TPEDSDPHFDPTALILEFTQDLCGVSTCCBTLGMADVGTVCDDPARSCAIVEDDGLQSAFTAA	360
QY	361 <td>HELGHVFMKLDNKRPCISLNGPLSTSRHWAPVMAHVDPPEPSPSCARFTFDLDNGY</td> <td>420</td>	HELGHVFMKLDNKRPCISLNGPLSTSRHWAPVMAHVDPPEPSPSCARFTFDLDNGY	420
Db	361 <td>HELGHVFMKLDNKRPCISLNGPLSTSRHWAPVMAHVDPPEPSPSCARFTFDLDNGY</td> <td>420</td>	HELGHVFMKLDNKRPCISLNGPLSTSRHWAPVMAHVDPPEPSPSCARFTFDLDNGY	420
QY	421 <td>GHCLLDKPEAPLHLPTVTGDKDD</td> <td>444</td>	GHCLLDKPEAPLHLPTVTGDKDD	444
Db	421 <td>GHCLLDKPEAPLHLPTVTGDKDD</td> <td>444</td>	GHCLLDKPEAPLHLPTVTGDKDD	444

DR EMBL: AB042272; BAB16474.1; -  
 DR EMBL: AB042271; BAB16473.1; -  
 DR EMBL: AB042273; BAB16475.1; -  
 DR InterPro: IPR001562; Disintegrin.  
 DR InterPro: IPR001590; Reptolysin.  
 DR InterPro: IPR000884; TSPL.  
 DR InterPro: IPR00130; Zn\_MpPtdase.  
 DR Pfam: PF00090; tspl\_1; 2.  
 DR Pfam: PF01421; Reptolysin.1.  
 DR SMART: SM00209; TSPL1.1.  
 DR PROSITE: PS50215; ADAM\_MEPPO.1.  
 DR PROSITE: PS50092; TSPL1.  
 DR PROSITE: PS00142; ZINC\_PROTEASE.1.  
 DR PROSITE: PS00427; DISINTEGRIN.1; FALSE NEG.  
 KW Hydrolase; Metalloprotease; zinc; Glycoprotein; zymogen;  
 KM Extracellular matrix.  
 FT PROPEP 1  
 FT CHAIN 5  
 FT METAL 6 630  
 FT ACT\_SITE 154 154  
 FT METAL 155 155  
 FT METAL 158 158  
 FT METAL 164 164  
 FT DOMAIN 233 303  
 FT DOMAIN 316 367  
 FT DOMAIN 368 478  
 FT DOMAIN 479 630  
 FT DOMAIN 40 45  
 FT CARBOHYD 96 96  
 FT CARBOHYD 474 474  
 SQ SEQUENCE 630 AA; 68384 MW; 63M428753167C7EF CRC64;

Query Match Best Local Similarity 48.3%; Score 1145; DB 1; Length 630;  
 Matches 217; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 208 RRARFASLSFEVETLVVADKMAFHGAGLKRLLTVMAAAAKAFKPSIRNPVSLVVT 267  
 DB 1 RRTFRFASLSFEVETLVVADKMAFHGAGLKRLLTVMAAAAKAFKPSIRNPVSLVVT 60  
 QY 268 RLVLGSGEEDPOGPRSAOGLRSCAMQRLNTPEDSDPHFTAILFTQDLCGVSTC 327  
 DB 61 RLVLGSGEEDPOGPRSAOGLRSCAMQRLNTPEDSDPHFTAILFTQDLCGVSTC 120  
 QY 328 DTLGMADVTCDDPARSCAIEDDGLQSAFTAAHELGHFVFMILNDNSKPCSLNGPLSTS 387  
 DB 121 DALGMAGVTCDDPARSCAIEDDGLQSAFTAAHELGHFVFMILNDNSKPCSLNGPLSTS 180  
 QY 388 RHVAPVMAHYDPEPSPSCSARFTITFDLNGYGHCLLDKPEAPLHLPTGDKYDD 444  
 DB 181 RHVAPVMAHYDPEPSPSCSARFTITFDLNGYGHCLLDKPEAPLHLPTGDKYDD 237

## RESULT 3

ATSL\_MOUSE STANDARD; PRT; 968 AA.  
 AC P97857; 054768; Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 41, Last annotation update)  
 DE 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE ADAMTS-1 Precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 1) (ADAM-TS1) (ADAM-TS1).  
 GN ADAMTS1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RP STRAIN=129/SVJ;  
 RX MEDLINE=98110583; PubMed=9441751;  
 RA Kuno K., Litzasa H., Ohno S., Matsushima K.;  
 RT "The exon/intron organization and chromosomal mapping of the mouse

RT ADAMTS-1 gene encoding an Adam family protein with TSP motifs.";  
 RL Genomics 46:466-471(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97150761; PubMed=8995297;  
 RA Kuno K., Kanada N., Nakashima E., Fujiki F., Ichimura F.,  
 RA Matsushima K.;  
 RT "Molecular cloning of a gene encoding a new type of metalloproteinase-  
 RT disintegrin family protein with thrombospondin motifs as an  
 RT inflammation associated gene.";  
 RL J. Biol. Chem. 272:556-562(1997).  
 RN [3]  
 RP CHARACTERIZATION, AND MUTAGENESIS OF GLU-403.  
 RX MEDLINE=99303657; PubMed=10373500.  
 RA Kuno K., Terasushima Y., Matsushima K.;  
 RT "ADAMTS-1 is an active metalloproteinase associated with the  
 RT extracellular matrix.";  
 RL J. Biol. Chem. 274:18821-18826(1999).  
 RN [4]  
 RP FUNCTION.  
 RX MEDLINE=20389568; PubMed=10930576;  
 RA Kuno K., Okada Y., Kawashima H., Nakamura H., Miyasaka M.,  
 RA Ohno H., Matsushima K.;  
 RT "ADAMTS-1 cleaves a cartilage proteoglycan, aggrecan.";  
 RL FEBS Lett. 478:241-245(2000).  
 RN [5]  
 RP FUNCTION, AND INDUCTION.  
 RX MEDLINE=20243757; PubMed=10781075;  
 RA Rodger R.L., Russell D.L., Espey L.L., Lydon J.P., O'Malley B.W.,  
 RA Richards J.S.;  
 RT "Progestosterone-regulated genes in the ovulation process: ADAMTS-1 and  
 RT cathepsin L proteases.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:4689-4694(2000).  
 CC -1- FUNCTION: CLEAVES AGGREGAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE  
 CC INVOLVED IN ITS TURNOVER. HAS ANGIOGENIC INHIBITOR ACTIVITY (BY  
 CC SIMILARITY). ACTIVE METALLOPROTEASE, WHICH MAY BE ASSOCIATED WITH  
 CC VARIOUS INFLAMMATORY PROCESSES AS WELL AS DEVELOPMENT OF CANCER  
 CC CACHEXIA. MAY PLAY A CRITICAL ROLE IN FOLLICULAR RUPTURE (BY  
 CC SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: CLEAVES AGGREGAN AT THE 1691-GLU-I-LEU-1692  
 CC SITE, WITHIN THE CHONDROITIN SULFATE ATTACHMENT DOMAIN.  
 CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: SECRETED. ASSOCIATED WITH THE EXTRACELLULAR  
 CC MATRIX.  
 CC -1- INDUCTION: INDUCED IN VITRO IN COLON ADENOCARCINOMA CELLS BY  
 CC INTERLEUKIN-1, OR IN VIVO IN KIDNEY AND HEART BY  
 CC LIPOPOLYSACCHARIDE. ALSO INDUCED BY LH STIMULATION IN GRANULOSA  
 CC CELLS OF PREOVULATORY FOLLICLES.  
 CC -1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT  
 CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.  
 CC -1- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2B.  
 CC -1- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 3 TSP TYPE-1 DOMAINS.  
 CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A  
 CC FRAMESHIFT IN POSITION 7.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: AB001735; BAA24501.1; ALT\_INT.  
 DR EMBL: D67076; BAA11088.1; ALT\_FRAME.  
 DR MEROPS: M12.222; -  
 DR MGD: MGI:109249; Adamts1.  
 DR InterPro: IPR001762; Disintegrin.  
 DR InterPro: IPR002870; Pep\_M12B\_propep.  
 DR InterPro: IPR001590; Reptolysin.  
 DR InterPro: IPR000884; TSPL.

Query Match	41.3%;	Score 979;	DB 1;	Length 968;
Best Local Similarity	46.8%;	Pred. No. 4.3e-64;		
Matches 212;	Conservative 66;	Mismatches 109;	Indels 66;	Gaps 14

QY	37	LLLLLLLS - LLLPSRLA - SPLREBEIVPEKLLNCGLVPG - SGTPARLLCRLQAGET	91
Db	35	MLLLLLLASITMLLCARGAHGHPTEDEDELIVPLSLEBA - PGHSTTTRL - RLDAFGQ	89
QY	92	LLLELEDDSSGVQEGELTYQYLQGAPELLGAE - - - - - PCTYLGTGTNGCPESVAS	14
Db	90	LHLKIQDPSGFLAPFETTLQDTGGRSP - - - - - GSEAQHLDTGDLAHCFYSGTNGDPGSA	14
QY	142	LHMDGALLGVLYQVRCAGELHLOPLEG - - - - - GTP - - - - - NSAGSGPAHILRRKSPASOG	19
Db	146	LSLCEG - VRGAFYIQGEEFFIQPARGVATELDRAVPEESSARQPHILRRRRGSC - G	20
QY	192	PMCNV - - - - - KAPLGSPSPRP - - - - - RAKRPASLSREVT	22
Db	204	AKCGMDEETLPTSDSPRESONTNRQMPVRPDTQDAGKPSGPGSIRKKRPVSSPRVET	26
QY	223	LVVAADDKMAAFHGGLKRYLLITVMAAAAKAKKHSSIRNPVSLYTRLVILGSGEGEQVG	28
Db	264	MLVADQSSADHVGSGGLKHYLLLTLESVAARFPRKHSIRNSISLVVAVKLLVLYEEKGGEVT	32
QY	283	PSAQTULRSPCAMOGRGLTPEDSRPHDFTALPTPRDLCVSTCDTFLGMADVTCYCPA	34
Db	324	SMAALTDLRPFNCNMOKQHNPSDRPHEHDTALPTTRDQCSHTCDTLIGMADVTCYCPDPS	38
QY	343	RSCAIVTEEDGLQSAFTAAHELGHVFNMLHDSKPCISLNGDLSTRHYMAFVMAHVPDEE	40
Db	384	RCSGVTEEDDGLQAAFTTAHEIGHVFNNPHDDAKRCASLNGVTGDS - HLMASMLSLDSHQ	44
QY	403	PWSPCSARFTIDFLDNGYTGCHLDKPRAPLHP	43
Db	443	PWSPCSAAMVTSTFLDNGEGELCKMKNPNTKLP	47

	RESULT 4		
ID	ATSL_RAT	STANDARD:	PRT: 967 AA.
AC	Q9WUO1: Q9ERL1;		
DT	16-OCT-2001 (Rel. 40, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1).		
GN	ADAMTS1..		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI_Taxid:10116;		
RA	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-Sprague-Dawley; TISSUE-Brain;		
RA	Liu X., Tu Y., Yin T., Johnstone E.M., Stephenson D.T., Clemens J.A.,		
RA	Little S.P.;		
RT	"Induction of a disintegrin and metalloprotease with the		
RT	thrombospondin type I motif (ADAMTS).";		
RL	Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.		
RL	[2]		
RP	SEQUENCE OF 18-967 FROM N.A.		
RC	STRAIN-Sprague-Dawley; TISSUE-Liver;		
RX	MEDLINE-20304009; PubMed-10847486;		
RA	Diamantis I., Luetchi M., Hoessl M., Relchen J.;		
RT	"Cloning of the rat ADAMTS-1 gene and its down regulation in		
RT	endothelial cells in cirrhotic rats.";		
RL	Layer 20:165-172(2000).		
CC	- FUNCTION: CLEAVES AGGRECAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE INVOLVED IN ITS TURNOVER. HAS ANGIOGENIC INHIBITOR ACTIVITY (BY SIMILARITY). ACTIVE METALLOPROTEASE, WHICH MAY BE ASSOCIATED WITH VARIOUS INFLAMMATORY PROCESSES AS WELL AS DEVELOPMENT OF CANCER CACHEKIA. MAY PLAY A CRITICAL ROLE IN FOLLICULAR RUPTURE (BY SIMILARITY).		
CC	- CATALYTIC ACTIVITY: CLEAVES AGGRECAN AT THE 1683-GLU- -LEU-1684 SITE, WITHIN THE CHONDROITIN SULFATE ATTACHMENT DOMAIN.		
CC	- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).		
CC	- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).		
CC	- INDUCTION: DOWN-REGULATED IN ENDOTHELIAL CELLS DERIVED FROM CIRRHOTIC LIVER.		
CC	- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.		
CC	- PFM: THE PRECURSOR IS CLEAVED BY A URIN ENDOPEPTIDASE (BY SIMILARITY).		
CC	- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.		
CC	- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.		
CC	- SIMILARITY: CONTAINS 3 TSP TYPE-1 DOMAINS.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send an email to license@sib-sib.ch).		
CC	or send an email to license@sib-sib.ch).		
CC	-----		
DR	EMBL: AF149118; AAC34012.1; -		
DR	EMBL: AF130446; AAG29823.1; -		
DR	MEROPS: M12.222; -		
DR	InterPro: IPR001762; Disintegrin.		
DR	InterPro: IPR002870; Rep_M12B_propep.		
DR	InterPro: IPR001590; Reptolysin.		
DR	InterPro: IPR000130; Zn_MTPepdse.		
DR	Pfam: PF00090; tsp_1; 6.		
DR	Pfam: PF01421; Reptolysin; 2.		
DR	Pfam: PF01563; pep_m12b_propep; 2.		
DR	SMART: SM00209; TSP1; 3.		
DR	PROSITE: PS50215; ADAM_MEPRO; 1.		
DR	PROSITE: PS00142; ZINC_PROTEASE; 1		



DR PROSITE: PS50092; TSP1; 2.  
 DR PROSITE: PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 KM Hydrolase: Metalloprotease; zinc: Signal: Glycoprotein; zymogen:  
 KW Repeat: Extracellular matrix; Heparin-binding.  
 FT SIGNAL 1 54  
 FT PROPEP 55 252  
 FT CHAIN 253 967  
 FT SITE 205 205  
 FT METAL 401 401  
 FT ACT\_SITE 402 402  
 FT METAL 405 405  
 FT METAL 411 411  
 FT DOMAIN 476 558  
 FT DOMAIN 616 724  
 FT DOMAIN 725 857  
 FT DOMAIN 858 907  
 FT DOMAIN 908 967  
 FT DOMAIN 194 198  
 FT CARBOHYD 547 547  
 FT CARBOHYD 720 720  
 FT CARBOHYD 764 764  
 FT CARBOHYD 782 782  
 FT CARBOHYD 945 945  
 FT CONFLICT 21 21  
 FT CONFLICT 26 31  
 FT CONFLICT 49 49  
 FT CONFLICT 72 72  
 FT CONFLICT 79 79  
 FT CONFLICT 249 249  
 FT CONFLICT 262 265  
 FT CONFLICT 607 607  
 FT CONFLICT 936 936  
 FT CONFLICT 962 962  
 SQ SEQUENCE 967 AA; 105705 MM; P93C864FDCB4C64C CRC64;  
 Query Match 41.2%; Score 977.5; DB 1; Length 967;  
 Best Local Similarity 46.5%; Pred. No. 5.5e-64;  
 Matches 210; Conservative 69; Mismatches 108; Indels 65; Gaps 13;

RESULT 5  
 ID ATSL\_HUMAN STANDARD: PRT: 967 AA.  
 AC Q9UHI8: Q9UHP80: Q9UHI83: Q9P2K0: Q9NSJ8;  
 DT 30-MAY-2000 (rel. 39, Created)  
 DT 16-OCT-2001 (rel. 40, Last sequence update)  
 DT 15-JUN-2002 (rel. 41, Last annotation update)  
 DE ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase  
 DE with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1) (METH-1).  
 GN ADAMTS1 OR METH1 OR KIA1346.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Casas C., Pritchard M.A., Estivill X., Arbones M.L.;  
 RT "Cloning, characterization and mapping on human chromosome 21 of the  
 RT orthologue of murine Adams-1.";  
 RT Submitted (JUL-1999) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A. AND FUNCTION.  
 RC Tissue-heart;  
 RX MEDLINE=99367466; PubMed=10438512;  
 RA Vazquez F., Hastings G., Ortega M.-A., Lane T.F., Oikemus S.,  
 RA Lombardo M., Iruela-Arispe M.L.;  
 RT "METH-1, a human ortholog of ADAMTS-1, and METH-2 are members of a new  
 RT family of proteins with angio-inhibitory activity.";  
 RL J. Biol. Chem. 274:23349-23357(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC Tissue-Endothelial cells;  
 RX MEDLINE=20247184; PubMed=10785405;  
 RA Glienke J., Schmitt A.O., Pilarsky C., Hinzmann B., Weiss B.,  
 RA Rosenthal A., Thierach K.H.;  
 RT "Differential gene expression by endothelial cells in distinct  
 RT angiogenic states.";  
 RL Eur. J. Biochem. 267:2820-2830(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC Tissue-Brain;  
 RX MEDLINE=20181126; PubMed=10718198;  
 RA Nagase T., Kikuno R., Ishikawa K.-I., Hirose M., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XVI.  
 RT The complete sequences of 150 new cDNA clones from brain which code  
 RT for large proteins in vitro.";  
 RL DNA Res. 7:65-73(2000).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20289799; PubMed=10830953;  
 RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,  
 RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Seda E.,  
 RA Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,  
 RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,  
 RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,  
 RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,  
 RA Shitani A., Sasaki T., Nagamine K., Miyasawa S., Antonarakis S.E.,  
 RA Minoshima S., Shimizu N., Nordstiek G., Hornischer K., Brandt P.,  
 RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecher H.,  
 RA Ramser J., Beck A., Klages S., Hennig S., Rlesselmann L., Dagand E.,  
 RA Wehmer J.S., Borzym K., Gardiner K., Nizetic D., Francis F.,  
 RA Lehrach H., Reinhardt R., Yaspo M.-L.;  
 RT "The DNA sequence of human chromosome 21.";  
 RL Nature 405:311-319(2000).  
 RN [6]  
 RP SEQUENCE OF 418-967 FROM N.A.  
 RC Tissue-Melanoma;  
 RA Blum H., Bauersachs S., Mewes H.-W., Weil B., Wiemann S.;  
 RT Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 CC - FUNCTION: CLEAVES AGGREGAN, A CARTILAGE PROTEOLYCAN, AND MAY BE  
 CC INVOLVED IN ITS TURNOVER (BY SIMILARITY). HAS ANGIOGENIC INHIBITOR  
 CC ACTIVITY. ACTIVE METALLOPROTEASE, WHICH MAY BE ASSOCIATED WITH  
 CC VARIOUS INFLAMMATORY PROCESSES AS WELL AS DEVELOPMENT OF CANCER

Query Match	39.9%	Score 946	DB 1	Length 967
Best Local Similarity	45.7%	Pred. No.	1,1e-61	
Matches 208	Conservative	62	Mismatches	119
			Indels	66
			Gaps	14

RESULT	6
ATSG_HUMAN	
ID	ATSG_HUMAN
AC	Q9UP79; Q9NZS0;
DT	16-OCT-2001 (Rel. 40, Created)
DT	16-OCT-2001 (Rel. 40, Last sequence update)
DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	ADAMTS-8 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 8) (ADAM-TS 8) (ADAM-TSB) (METH-2) (METH-8).
DE	
GN	ADAMTS8 OR METH8.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_Taxid:9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Lung;
RX	MEDLINE=93j67466; PubMed=10438512;
RA	Vazquez F., Hastings G., Ortega M.-A., Lane T.F., Olkemus S., Lombardo M., Iruela-Arispe M.L.;
RT	"METH-1, a human ortholog of ADAMTS-1 and METH-2 are members of a new family of proteins with angio-inhibitory activity.";
RL	J. Biol. Chem. 274:23349-23357(1999). [2]
RN	
RP	SEQUENCE OF 195-440 FROM N.A.
RX	MEDLINE=20079168; PubMed=10610729;
RA	Georgiadis K.E., Hirohata S., Seidin M.F., Apte S.S.;
RT	"ADAM-TS8, a novel metalloprotease of the ADAM-TS family located on mouse chromosome 9 and human chromosome 11.";
RL	Genomics 62:312-315(1999).
CC	-1- FUNCTION: HAS ANTI-ANGIOGENIC PROPERTIES.
CC	-1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC	-1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).
CC	-1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN ADULT AND FETAL LUNG, LOWER EXPRESSION IN BRAIN, PLACENTA, HEART, STOMACH AND FETAL BRAIN AND KIDNEY.
CC	-1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC	



DR Pfam: PF00090; tsp-1; 2.  
 DR Pfam: PF01421; Reprolysin; 1.  
 DR SMART: SM00209; TSP1; 2.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 DR PROSITE: PS500215; ADAM\_MEPRO; 1.  
 DR PROSITE: PS50092; TSP1; 1.  
 DR PROSITE: PS00427; DISINTEGRIN-1; FALSE\_NEG.  
 DR Hydrolase: Metalloprotease; zinc; signal; Glycoprotein; zymogen;  
 KW Repeat: Extracellular matrix; Heparin-binding.  
 FT SIGNAL 1 28  
 FT PROPEP 29 238  
 FT CHAIN 229 905  
 FT METAL 378 378  
 FT ACT\_SITE 379 379  
 FT METAL 382 382  
 FT METAL 388 388  
 FT DOMAIN 453 541  
 FT DOMAIN 542 598  
 FT DOMAIN 599 705  
 FT DOMAIN 706 847  
 FT CARBOHYD 848 905  
 FT CARBOHYD 415 415  
 FT CARBOHYD 480 480  
 FT CARBOHYD 506 506  
 FT CARBOHYD 615 615  
 SQ SEQUENCE 905 AA; 98879 MW; 124D4132B3A0CAE CRC64;

Query Match 35.7%; Score 846.5; DB 1; Length 905;  
 Best Local Similarity 41.2%; Pred. No. 26-54; Indels 85; Gaps 15;  
 Matches 192; Conservative 65; Mismatches 124;

36 WLLLLLLASLPASRLASPLP-----REEIVPEPKLNGSVLPSSGTPARLLCRLQAF 88  
 10 WPLLLLLLPPLPPVPCAGAPGCGAASLVPTRLPGSA-----SELAFSLNAF 62  
 89 GELLLELLEODSVOYEGVLQYLGAPLGLGAE----GYLTGTINDPESVASLH-- 143  
 63 GQGEVLRLLPDSFELAPERKIERLGSSAAAGG-EPGLRGCFESGTVNRESLAMSCV 121  
 144 --WDGGLLVGLVYRGAEHLQPLGEG---TPN--SAGPGAHILRRKSP----- 186  
 122 AGSSGSFLA-----GEFTIQPGAGDLDOPHRLOKNGPQ---RRDPGLAAAEVPP 173  
 187 -----ASGCG-----PMCVKAPLGSPPRRRAKRF 213  
 174 LPQGLEMEVEMGCGQOERSNEDERKQDEKGLKETEDSRKPPFGSKT---RSKRF 229  
 214 ASLSREVELLVADDMAPFHGAGLRYLLTWAAAKAFKHPSTINPVSLLVTRVLIG 273  
 230 VSEARVEVLVADASMAAFYCTDLONLITVWSMAARIYKHPSTINSVNLVVKVLVE 289  
 274 SGEQPGVPSAAQTLRSFCAMQRCGLNTPEDSDPHFDALIFTRDGLGV-STCDTGLM 332  
 290 KEMWGEVSDNGCLTRLNCSMQRNKRKSDRPHETVTRAILFTKRNFCGKGCQDTLGM 349  
 333 ADVGVTCDPARSCAIVEDGLQSAFTAAHELGHVENMLHDNSKPCISLNGPLSTSRHVA 392  
 350 ADVGTICDDKSCSYIKDGLQAAVTLAHELGHVLSMPDDSKPCVRLGPMG-KYHMAA 408  
 393 PVMANVDPPEPSPCARITDFDLNGYGHCLLDKPEALHLPVYG 438  
 409 PFFIHVNTKTLPMSPCAVYLTLLDDGHCDCLLDAPTSLPLP-TG 453

RESULT 8  
 ATSS\_MOUSE STANDARD; PRT; 930 AA.  
 ID ATSS\_MOUSE  
 DT 16-OCT-2001 (Rel. 40; Created)  
 DT 16-OCT-2001 (Rel. 40; Last sequence update)  
 DT 15-JUN-2002 (Rel. 41; Last annotation update)  
 DE ADAMTS-5 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 5) (ADAM-TS 5) (Aggrecanase-2)

DE (ADMP-2) (Implantin).  
 GN ADAMTS5.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-99395124; PubMed-10464288;  
 RA Hurskainen T.L., Hitchaba S., Seldin M.F., Apte S.S.:  
 \*ADAM-TS5, ADAM-TS6, and ADAM-TS7, novel members of a new family of  
 zinc metalloproteases.\*  
 RT J. Biol. Chem. 274:25555-25563(1999).  
 RL  
 CC -1- FUNCTION: CLEAVES AGGREGAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE  
 INVOLVED IN ITS TURNOVER. MAY PLAY AN IMPORTANT ROLE IN THE  
 DESTRUCTION OF AGGREGAN IN ARTHRITIC DISEASES. MAY PLAY A ROLE IN  
 PROTEOLYTIC PROCESSING MOSTLY DURING THE PERI-IMPLANTATION PERIOD.  
 CC -1- CATALYTIC ACTIVITY: Cleaves aggrecan at the 392-Glu-Ala-393  
 site.  
 CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular  
 matrix (By similarity).  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED SPECIFICALLY IN THE PERI-  
 IMPLANTATION PERIOD IN EMBRYO AND TROPHOBLAST AND AT LOW OR  
 UNDETECTABLE LEVEL THEREAFTER.  
 CC -1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT  
 FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.  
 CC -1- PTM: THE PRECURSOR IS CLEAVED BY A PURIN ENDOPEPTIDASE (BY  
 SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.  
 CC -1- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 2 TSP TYPE-1 DOMAINS.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 -----  
 CC EMBL: AF140673; AAD56356.1; -.  
 DR MEROPS: M12.225; -.  
 DR MGD: MGI:1346321; Adamts5.  
 DR InterPro: IPR001762; Disintegrin.  
 DR InterPro: IPR002870; Pep\_M12B\_propep.  
 DR InterPro: IPR001590; Reprolysin.  
 DR InterPro: IPR000884; TSP1.  
 DR Pfam: PF00090; tsp-1; 2.  
 DR Pfam: PF01421; Reprolysin; 1.  
 DR SMART: PS001562; Pep\_M12B\_propep; 1.  
 DR SMART: SM00209; TSP1; 2.  
 DR PROSITE: PS50215; ADAM\_MEPRO; 1.  
 DR PROSITE: PS50092; TSP1; 1.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 DR PROSITE: PS00427; DISINTEGRIN-1; FALSE\_NEG.  
 KW Hydrolase: Metalloprotease; zinc; signal; Glycoprotein; zymogen;  
 KW Repeat: Extracellular matrix.  
 FT SIGNAL 1 21  
 FT PROPEP 22 261  
 FT CHAIN 262 930  
 FT SITE 209 209  
 FT METAL 410 410  
 FT ACT\_SITE 411 411  
 FT METAL 414 414  
 FT METAL 420 420  
 FT METAL 485 566  
 FT DOMAIN 567 623  
 FT DOMAIN 624 731  
 FT DOMAIN 732 874  
 FT DOMAIN 875 930  
 FT DOMAIN 41 46

```

FT DOMAIN 257 261 POLY-ARG. (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 498 498 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 728 728 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 802 802 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 807 807 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 930 AA: 101780 MW: 84DE84B26170D4DC CRC64;

Query Match 27.9%; Score 661; DB 1; Length 930;
Best Local Similarity 34.6%; Pred. No. 8.1e-41;
Matches 170; Conservative 69; Mismatches 160; Indels 92; Gaps 14;

QY 31 LSW-LVWLLLLLSL-PSKRLASLP-KEELY 62
DB 3 LEWASLLLSLSASCLSLAADSPAAPADRTROPOAAAAAEPDPQGEETREHGL 62
QY 63 FP-EKINGSVL-PCSGTPARLLCLQAFGETLLELLEPDSQVYEGTLVQL 112
DB 63 QPLAGCRSGGIVNHIDOLYSGGKVGTL-VYAGGRFLDLERDVTGAAQ-SIYTA 118
QY 113 GQAPELLGAEPTVLTGTINGDPESVASLHWDGALLGVLYRGAELHLOPLEGG- 168
DB 119 GGGLSASSGHRCHFCYRGVDSFSLAVFDLCG-LDGFVAKHARTYLRKLGSMAE 177
QY 169 -----TPNSAG-----PGAH-LRKSPASQ 190
DB 178 YERIVGDSRRILHVNREGFSEALPPRASCETPASPSPGDESPSVSRSRRSALAPQ 237
QY 191 GPMCVKAPGSPDRP-RRAKRFASLSRFVETLVVADDDKMAFHAGLKRKYLTVMAA 248
DB 238 LLDHAFSPSGMAGQOTWRRRRRSISRAROVYELLVADSSMARWYGRLOHLYLLAST 297
QY 249 AAKAFKPSIRNPVSLVYTVLTVLGSGEQPGVPSAOTLRSFCAMORGLTTPEDSPD 308
DB 298 ANRLSHASIEHRIKAVKVVVLTDKPTSLVSKNAATTLKPFCKWQHQNQLDDHE 357
QY 369 HEDTILTRDOLCVSTCDPLGMADVCTCDPARSCAIVEDDGLQSAFTAHETGHVFN 368
DB 368 HYDAILTRERDLCGHSDTLGMADVCTGSPERSCAVIEDGLHAFTVAHETGHLG 417
QY 369 MLHDSKPCISLNGPLSTSRHYMAVMAHVDPEPSPSCAFITFDLNGYGHCLDKP 428
DB 418 LSHDSKPC-EENFGTDEKRLMSSILTSIDASKPMSKCTSTITFEFLDNGHCLDLDP 476
QY 429 E----APLHLP 435
DB 477 RKQILGPELPL 487

RESULT 9
ATTS_HUMAN STANDARD: PRF: 930 AA.
ID ATTS_HUMAN Q9UNA0: Q9URP2;
AC Q9UNA0: Q9URP2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE ADAMTS-3 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 5) (ADAM-TS 5) (ADAM-TS5) (Aggrecanase-2) (ADMP-2) (ADAM-TS 11).
GN ADAMTS5 OR ADMP2 OR ADAMTS11.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_taxid:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Liver;
RX MEDLINE-99367476; PubMed-10438522;
RA Abbaszade I., Liu R.-Q., Yang F., Rosenfeld S.A., Ross O.H., Link J.R., Ellis D.M., Tortorella M.D., Pratta M.A., Hollis J.M., Wynn R., Duke J.L., George H.J., Hillman P.C. Jr., Murphy K., Mswall B.H., Copeland R.A., Decicco C.P., Bruckner R., Nagase H., Ito Y., Newton R.C., Magolda R.L., Trzaskos J.M., Hollis G.F., Arner E.C., Burn T.C.;

```

```

RT "Cloning and characterization of ADAMTS11, an aggrecanase from the
RT ADAMTS family.";
RT J. Biol. Chem. 274:23443-23450(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-20289799; PubMed-10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Matanabe H., Yada T., Park H.-S., Toyoda A., Ishii K., Toloki Y., Choi D.-K., Soeda E., Onki M., Takegi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D., Reichwald K., Rump A., Schillnabel M., Schudy A., Zimmermann W., Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S., Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E., Minoshima S., Shimizu N., Nordstiek G., Hornischer K., Brandt P., Schaefer M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecher H., Ransner J., Beck A., Klages S., Hennig S., Rieselmann L., Dagand E., Wehmeyer S., Botzrym K., Gardiner K., Nizetic D., Francis F., Lehnach H., Reinhardt R., Yaspo M.-L.;
RA "The DNA sequence of human chromosome 21.";
RT Nature 405:311-319(2000).
RN [3]
RP SEQUENCE OF 413-930 FROM N.A.
RC TISSUE-Fetal brain;
RX MEDLINE-99395124; PubMed-10464288;
RA Hurskainen T.L., Hirohata S., Seidin M.F., Apte S.S.;
RT "ADAM-TS5, ADAM-TS6, and ADAM-TS7, novel members of a new family of zinc metalloproteinases.";
RT J. Biol. Chem. 274:25555-25563(1999)
CC -1- FUNCTION: CLEAVES AGGREGAN, A CARTILAGE PROTEOLYCAN, AND MAY BE INVOLVED IN ITS TURNOVER. MAY PLAY AN IMPORTANT ROLE IN THE DESTRUCTION OF AGGREGAN IN ARTHRITIC DISEASES. MAY PLAY A ROLE IN PROTEOLYTIC PROCESSING MOSTLY DURING THE PERI-IMPLANTATION PERIOD.
CC -1- CATALYTIC ACTIVITY: Cleaves aggrecan at the 392-Glu1-Ala-393 site.
CC -1- COPOLYMER: BINDS 1 ZINC ION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (by similarity).
CC -1- TISSUE SPECIFICITY: EXPRESSED AT LOW LEVEL, PRIMARILY IN PLACENTA BUT ALSO IN OTHER TISSUES, SUCH AS HEART AND BRAIN, AND ALSO CERVIX, UTERUS, BLADDER, ESOPHAGUS, RIB CARTILAGE, CHONDROBLASTOMA, FIBROUS TISSUE AND JOINT CAPSULE FROM AN ARTHRITIC PATIENT.
CC -1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC -1- PFM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC -1- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 TSP TYPE-1 DOMAINS.
CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
CC CC
DR EMBL: AF142099; AAD9577.1; -
DR EMBL: AP001698; BAA95504.1; -
DR EMBL: AP001697; BAA95503.1; -
DR EMBL: AF141293; AAF02493.1; -
DR HSSP: Q9PW35; 1BDU.
DR MEROPS: M12.225; -.
DR Gene: HGNC:221; ADAMTS5.
DR MIM: 605007;
DR InterPro: IPR001762; Disintegrin.
DR InterPro: IPR002870; Pep_M12B_propep.
DR InterPro: IPR001590; Reprolysin.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR000130; Zn_MTPepidase.
DR Pfam: PF00090; tsp_1; 2.
DR Pfam: PF01421; Reprolysin; 1.

```

DR Pfam: PF01562; Pep\_M12B\_propep; 1.  
 DR SMART; SM00209; TSP1; 2.  
 DR PROSITE; PS00215; ADAM\_MERPRO; 1. FALSE\_NEG.  
 DR PROSITE; PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 DR PROSITE; PS50092; TSP1; 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 DR HydroLase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;  
 KW Repeat: Extracellular matrix.  
 FT SIGNAL 1 16  
 FT PROPEP 17 261  
 FT CHAIN 262 930  
 FT SITE 209 209  
 FT METAL 410 410  
 FT ACT\_SITE 411 411  
 FT METAL 414 414  
 FT METAL 420 420  
 FT METAL 455 456  
 FT DOMAIN 567 623  
 FT DOMAIN 624 732  
 FT DOMAIN 732 874  
 FT DOMAIN 875 930  
 FT DOMAIN 930 930  
 FT DOMAIN 930 930  
 FT CARBOHYD 257 261  
 FT CARBOHYD 498 498  
 FT CARBOHYD 728 728  
 FT CARBOHYD 802 802  
 FT CARBOHYD 807 807  
 FT CONFLICT 138 138  
 FT CONFLICT 614 614  
 FT CONFLICT 692 692  
 SQ SEQUENCE 930 AA; 101715 MW; B64281502F28193B CnC64;  
 Query Match 27.5%; Score 651.5; DB 1; Length 930;  
 Best Local Similarity 33.7%; Pred. No. 4e-40;  
 Matches 168; Conservative 64; Mismatches 156; Indels 111; Gaps 15;

RESULT 10  
 ID AT59\_HUMAN STANDARD; PRT; 1629 AA.  
 AC 09P2N4; 09NR29;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE ADAMTS-9 precursor (BC 3.4.24.-) (A disintegrin and metalloproteinase  
 DE with thrombospondin motifs 9) (ADAM-TS 9) (ADAM-TS9).  
 GN ADAMTS9 OR KIA1312.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;  
 [1]  
 RP SEQUENCE FROM N.A. (SHORT ISOFORM).  
 RC TISSUE=Fetal.  
 RX MEDLINE=20396138; PubMed=10936055;  
 RA Clark M.E., Kelnner G.S., Turbeville L.A., Boyer A., Arden K.A.,  
 RA Maki R.A.;  
 RT \*ADAMTS 9, a novel member of the ADAM-TS/metalloprotein gene  
 RT family.\*;  
 RL Genomics 67:343-350(2000).  
 RN [2]  
 RP SEQUENCE OF 159-1629 FROM N.A. (LONG ISOFORM).  
 RC TISSUE=Brain;  
 RX MEDLINE=20181126; PubMed=10718198;  
 RA Nagase T., Kikuno R., Ishikawa K.-I., Hirosewa M., Ohara O.;  
 RT \*Prediction of the coding sequences of unidentified human genes. XVI.  
 RT The complete sequences of 150 new cDNA clones from brain which code  
 RT for large proteins in vitro.\*;  
 RL DNA Res. 7:65-73(2000).  
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular  
 CC matrix (By similarity).  
 CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: Highly expressed in all fetal tissues.  
 CC -1- EXPRESSED SLIGHTLY IN ADULT OVARY, PANCREAS, HEART, KIDNEY, LUNG,  
 CC PLACENTA. ALSO DETECTED IN SPINAL CORD AND BRAIN. NOT DETECTED IN  
 CC COLON, SMALL INTESTINE, TESTIS, LIVER, SKELETAL MUSCLE, SPLEEN OR  
 CC THYMUS.  
 CC -1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT  
 CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX (BY  
 CC SIMILARITY).  
 CC -1- PTM: THE PRECURSOR IS CLEAVED BY A FUZIN ENDOPEPTIDASE (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.  
 CC -1- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 11 TSP TYPE-1 DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: AF261918; AAF89106.1; -;  
 DR EMBL: AB037733; BAA92550.1; -;  
 DR HSSP: P15167; IATL.  
 DR MEROPS: M12.021; -;  
 DR Genem: HGNC:13202; ADAMTS9.  
 DR MIM: 605421; -;  
 DR InterPro: IPR001762; Disintegrin.  
 DR InterPro: IPR002870; Pep\_M12B\_propep.  
 DR InterPro: IPR001590; Reprolysin.  
 DR InterPro: IPR000884; TSP1.  
 DR InterPro: IPR000130; Zn\_MTPeptide.  
 DR Pfam: PF00090; tsp\_1; 11.  
 DR Pfam: PF01421; Reprolysin; 1.  
 DR Pfam: PF01562; Pep\_M12B\_propep; 1.

```

Db 350 EODGDSISFNQTLKNCQMOHNSKNSGGI---HHDTAVLLTRQDLCRAHDKDCTGLA 406
Oy 334 DVGIVCDPASCALVEDGLOSAFTAHAEHLGVNRLMLHDSKSPCLSLNGPLSTSRHWAP 393
Db 407 ELGRTIDPYRSCSISEDGLSTAFIAHEHLGVNMPMDNNKC--KEEGVKSPOHWAP 464
Oy 394 VMAHNDPEEWPSCARFITFDLNGVGHCLLDKREA-PLHLPPV 436
Db 465 TLNFYTNFPMWMSKSRKTYTERLDTGYEGCLLNEPESRPYLPV 508

RESULT 11
AT54_BOVIN STANDARD: PRT: 245 AA.
ID_AT54_BOVIN
AC O9PT93;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE ADAMTS-4 (EC 3.4.24.-) (A disintegrin and metalloproteinase with
thrombospondin motifs 4) (ADAM-TS 4) (ADAM-TS4) (Aggrecanase 1) (ADMP-
1) (Fragments).
DE ADAMTS4.
GN Bos taurus (bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
OX NCBI_TaxID:9913;
[1]
RN RP SEQUENCE OF 18-227 FROM N.A.
RX MEDLINE:99333677; PubMed-10403768;
RA Flannery C.R., Little C.B., Hughes C.E., Caterson B.;
RT "Expression of ADAMTS homologues in articular cartilage.";
RL Biochem. Biophys. Res. Commun. 260:318-322(1999).
[2]
RN RP SEQUENCE OF 18-227 FROM N.A.
RX MEDLINE:20092827; PubMed-10625599;
RA Curtis C.L., Hughes C.E., Flannery C.R., Little C.B., Hatwood J.L.,
RA Caterson B.;
RT "n-3 fatty acids specifically modulate catabolic factors involved in
articular cartilage degradation.";
RL J. Biol. Chem. 275:721-724(2000).
[3]
RN RP SEQUENCE OF 1-27: 228-233 AND 234-245.
RC TISSUE-Cartilage;
RX MEDLINE:99286303; PubMed-10356395;
RA Tortorella M.D., Burn T.C., Pratta M.A., Abbazade I., Hollis J.M.,
RA Liu R., Rosenfeld S.A., Copeland R.A., Decicco C.P., Wynn R.,
RA Rockwell A., Yang F., Duke J.L., Solomon K., George H., Bruckner R.,
RA Nagase H., Itoh Y., Ellis D.M., Ross H., Wiswall B.H., Murphy K.,
RA Hillman J.C. Jr., Hollis G.F., Newton R.C., Magolda R.L.,
RA Tzikos J.M., Aher E.C.;
RT "Purification and cloning of aggrecanase-1: a member of the ADAMTS
family of proteases.";
RL Science 284:1664-1666(1999).
CC -1- FUNCTION: CLEAVES AGGECAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE
INVOLVED IN ITS TURNOVER. MAY PLAY AN IMPORTANT ROLE IN THE
DESTRUCTION OF AGGECAN IN ARTHRITIC DISEASES.
CC -1- CATALYTIC ACTIVITY: Cleaves aggrecan at the 392-Glu-I-Ala-393
site.
CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
matrix (by similarity).
CC -1- INDUCTION: BY INTERLEUKIN-1.
CC -1- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY
SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M17H.
CC -1- CAUTION: HAS SOMETIMES BEEN REFERRED TO ADAMTS2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: AF192770; AAF07176.1; -

DR HSSP: P34179; AAF07176.1; -

DR InterPro: IPR001590; Reprolysin.

DR InterPro: IPR001590; Zn\_MTPeptide.

DR PROSITE: PS00142; ZINC\_PROTEASE; 1.

DR PROSITE: PS50215; ADAM\_MEPPO; 1.

KM Hydrolase; Metalloprotease; Zinc; Glycoprotein; Extracellular matrix.

FT NON\_TER 1

FT NON\_CONS 1

FT METAL 27

FT ACT\_SITE 57

FT METAL 58

FT METAL 61

FT METAL 67

FT NON\_CONS 227

FT NON\_CONS 233

FT NON\_TER 245

SO SEQUENCE 245 AA; 26391 MW; 7C687968CAEB431D CRC64;

Query Match

Best Local Similarity 56.0%; Pred. No. 1.7e-39; Length 245;

Matches 130; Conservative 2; Mismatches 8; Indels 92; Gaps 1;

QY 213 FASLSRFVETLVVADDKMAAFHCGAGLKRLLTVMAAFAKHPKIRNPVSLVTVTLV 272

DB 1 FASLSRFVETLVVADDKMAAFHCGAGLKRLLTVMAAFAKHPKIRNPVSLVTVTLV 27

QY 273 GSGEERPOVGSAAQTLRSCAMQRLNTPEDSDPHFTAILFTMODLCGVSTCDLGM 332

DB 28 -----

QY 333 ADVGTCDPARSAICVIEDDGLGSAFTAAHGLHVFNMHLDNSKPCISLNGPLSTRHVA 392

DB 29 ADVGTCDPARSAICVIEDDGLGSAFTAAHGLHVFNMHLDNSKPCISLNGPLSTRHVA 88

QY 393 PVAHVDPPEPMSPCSAFRTTDFLNGVGHCLDKPEAPLHLPVTDYKDD 444

DB 89 PVAHVDPPEPMSPCSAFRTTDFLNGVGHCLDKPEAPLHLPVTDYKDD 140

RESULT 12

AT57\_HUMAN

ID AT57\_HUMAN STANDARD: PRT: 997 AA.

AC Q9URP4;

DT 16-OCT-2001 (Rel. 40; Created)

DT 16-OCT-2001 (Rel. 40; Last sequence update)

DT 15-JUN-2002 (Rel. 41; Last annotation update)

DE ADAMTS-7 precursor (RC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 7) (ADAM-TS 7) (ADAM-TS7).

GN ADAMTS7.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Carnivora; Hominoidea; Homo.

OX NCBI\_TaxID=9606;

RP SEQUENCE FROM N.A.

RX MEDLINE-99395124; PubMed-10464288;

RA Hurskainen T.L., Hirohata S., Seidlin M.F., Apte S.S.;

RT "ADAM-TS5, ADAM-TS6, and ADAM-TS7, Novel Members of a New Family of Zinc Metalloproteases."

RL J. Biol. Chem. 274:25555-25563(1999).

CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).

CC -1- TISSUE SPECIFICITY: EXPRESSED IN HEART, BRAIN, PLACENTA, LUNG, LIVER, SKELETAL MUSCLE, KIDNEY, PANCREAS.

CC -1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.

CC -1- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.

CC -1- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.

CC -1- SIMILARITY: CONTAINS 2 TSP TYPE-1 DOMAINS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: AF140675; AAD56358.1; -

DR HSSP: P15167; IATL.

DR MEROPS: M12.231; -

DR Genew: HGNC:223; ADAMTS7.

DR MIM: 605009; -

DR InterPro: IPR001762; Disintegrin.

DR InterPro: IPR002870; Pep\_M12B\_propep.

DR InterPro: IPR001590; Reprolysin.

DR InterPro: IPR000884; TSP1.

DR InterPro: IPR000130; Zn\_MTPeptide.

DR Pfam: PR00090; tsp\_1; 1.

DR Pfam: PR01421; Reprolysin; 1.

DR Pfam: PF01562; Pep\_M12B\_propep; 1.

DR SMART: SM00209; TSP1; 2.

DR PROSITE: PS00142; ZINC\_PROTEASE; 1.

DR PROSITE: PS50215; ADAM\_MEPPO; 1.

DR PROSITE: PS00427; DISINTEGRIN\_1; FALSE\_NEG.

KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen; Repeat; Extracellular matrix.

KW SIGNAL

FT SIGNAL 1

FT PROPEP 27

FT CHAIN 28

FT SITE 233

FT SITE 204

FT METAL 388

FT ACT\_SITE 389

FT METAL 392

FT METAL 398

FT METAL 462

FT DOMAIN 537

FT DOMAIN 538

FT DOMAIN 544

FT DOMAIN 595

FT DOMAIN 697

FT DOMAIN 698

FT DOMAIN 915

FT CARBOHYD 94

FT CARBOHYD 693

FT CARBOHYD 778

SO SEQUENCE 997 AA; 109694 MW; 6587044ED02FC104 CRC64;

Query Match

Best Local Similarity 34.4%; Pred. No. 1.6e-31; Length 997;

Matches 155; Conservative 57; Mismatches 165; Indels 73; Gaps 17;

QY 38 LLLLSLSPARSLAPLPRPREE-----ELVPEKLN--GSVLPGSTPARLLCR----- 84

DB 17 LLLLSLSPARSLAPLPRPREE-----ELVPEKLN--GSVLPGSTPARLLCR----- 84

QY 85 -----LQAFGFTLLLEEDSGVVEGLTVQY-----LQAFELLGAEFGTYLTGTI 132

DB 77 DAPAFELQYRGHLELFNLTANOHLAPGVSETRRGGLGRAH--IRAHTRACHILGVE 134

QY 133 NGDPESVASLHMDG-----ALLGVLYRGAEHLHLPLEGCTPPNSAGCGAHIL--RR 183

DB 135 Q-DPE-----LEGGLAATSACDGLKGVQLNSDEVFIPLD-SAPARGAHQPHVYKR 186

QY 184 KSPA-----SGQGPWCNVKAPLGSFS-----PPRRA-KRFAISLRFVET 222

DB 187 QAPERLARGDSASPTCCGVYPPLESRRERREORQORQRRRLRLHQRYSKKEKWT 246

QY 223 LVVADDKMAAFHCGAGLKRLLTVMAAFAKHPKIRNPVSLVTVTLVILSGEGEPQV 281

DB 247 LVVADDKMAAFHCGAGLKRLLTVMAAFAKHPKIRNPVSLVTVTLVILSGEGEPQV 306



QY 282 GPSAAQTLKSPCAMORGLNTPEDSPDHFTALFTRODLCGVST--CDTLGMADVGYC 339  
 DB 307 THHADNTLSPCKMKSIMKGAHHLHHDFTAILTRKLCACAMNRCETLGLSHVAGMC 366  
 QY 340 DPARSCAIVEDDGLSAFTAHNELGHVFNMLHDNS-KPCISLNGPLSTRHVAAPMAHV 398  
 DB 367 QHRSCSINEDGLPLAFVAHGLHSGFIOHDSGNDG----EPGKRPFIMSPQLYD 422  
 QY 399 DPEPSPCSARFTDPLDNGYGHCLDKP 428  
 DB 423 AAPLTWRSRCSROYITRFLDRGWGLCLDDPP 452

RESULT 13  
 AT12\_HUMAN STANDARD: PRT: 1593 AA.  
 ID AT12\_HUMAN  
 AC P58397;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE ADAMTS-12 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 12) (ADAM-TS 12) (ADAM-TS12)  
 GN ADAMTS12.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Fetal lung;  
 RX MEDLINE=21264577; PubMed=11279086;  
 RA Cal S., Argueelles J.M., Fernandez P.L., Lopez-Otin C.;  
 RT Identification, characterization, and intracellular processing of ADAM-TS12, a novel human disintegrin with a complex structural organization involving multiple thrombospondin-1 repeats.\*;  
 RL J. Biol. Chem. 276:17932-17940(2001).  
 CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (by similarity).  
 CC -1- TISSUE SPECIFICITY: Expressed exclusively in fetal lung. Is widely expressed in gastric carcinomas and in cancer cells of diverse origin.  
 CC -1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE 1 DOMAINS ARE IMPORTANT FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX (BY SIMILARITY).  
 CC -1- PTM: THE PRECURSOR IS CLEAVED BY A URIN ENDOPEPTIDASE.  
 CC -1- PTM: IS SUBJECTED TO AN INTRACELLULAR MATURATION PROCESS LEADING TO A FRAGMENT CONTAINING THE N-TERMINAL REGION INCLUDING THE METALLOPROTEINASE, DISINTEGRIN-LIKE, CYS-RICH AND TS-1 DOMAINS AND THE C-TERMINAL FRAGMENT CONTAINING THE SPACER 2 AND THE FOUR TS-1 DOMAINS.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.  
 CC -1- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 8 TSP TYPE-1 DOMAINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL: AJ250725; CAC20419.1;  
 CC Genew; HGNC:14605; ADAMTS12.  
 DR MIM: 606184;  
 DR InterPro: IPR001762; Disintegrin.  
 DR InterPro: IPR002870; Pep\_M12B\_propep.  
 DR InterPro: IPR001590; Repolysin.  
 DR InterPro: IPR000864; TSP1.  
 DR InterPro: IPR000130; zn\_MTPeplase.  
 DR Pfam: PF00090; tsp\_1; 6.

DR Pfam: PF01421; Repolysin; 1.  
 DR Pfam: PF01562; Pep\_M12B\_propep; 1.  
 DR SMART: SM00209; TSP1; 8.  
 DR PROSITE: PS50215; ADAM\_MEPRO; 1.  
 DR PROSITE: PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 DR PROSITE: PS50092; TSP1; 2.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 KW Hydrolase; Metalloproteinase; Zinc; Signal; Glycoprotein; Zymogen;  
 KW Repeat; Extracellular matrix.  
 FT SIGNAL 1 25  
 FT PROPEP 26 240  
 FT CHAIN 241 1593  
 FT DOMAIN 465 544  
 FT DOMAIN 545 596  
 FT DOMAIN 597 700  
 FT DOMAIN 701 826  
 FT DOMAIN 827 881  
 FT DOMAIN 886 943  
 FT DOMAIN 947 995  
 FT DOMAIN 996 1315  
 FT DOMAIN 1316 1364  
 FT DOMAIN 1367 1423  
 FT DOMAIN 1426 1471  
 FT DOMAIN 1471 1471  
 FT DOMAIN 302 305  
 FT SITE 208 208  
 FT SITE 392 392  
 FT ACT\_SITE 393 393  
 FT METAL 396 396  
 FT METAL 402 402  
 FT CARBOHYD 105 105  
 FT CARBOHYD 125 125  
 FT CARBOHYD 215 215  
 FT CARBOHYD 485 485  
 FT CARBOHYD 685 685  
 FT CARBOHYD 790 790  
 FT CARBOHYD 951 951  
 FT CARBOHYD 1104 1104  
 FT CARBOHYD 1275 1275  
 FT CARBOHYD 1300 1300  
 FT CARBOHYD 1320 1320  
 FT CARBOHYD 1371 1371  
 FT CARBOHYD 1378 1378  
 FT CARBOHYD 1503 1503  
 SO SEQUENCE 1593 AA; 177545 MM; 07F9F48B63BD83A3 CRC64;

Query Match 20.8%; Score 493.5; DB 1; Length 1593;  
 Best local similarity 35.0%; Pred. No. 2,9e-28;  
 Matches 117; Conservative 58; Mismatches 132; Indels 27; Gaps 9;

QY 117 ELUGAEPGYLTGT--KNDPESVASLHMDGALLGVLDYRGAEHLHLPLEGCTRSAG 174  
 DB 130 KMASSAPLCHLSGTVLQOQTRVGTALASCHG-LTGFEQLPHGDFIEYK-KHPLVEG 187  
 QY 175 GPCAHIL-RRKSPASOGPCWKNKAPLGSPS-----PRPRKAFASLRP 219  
 DB 188 GYRPHIVYRKQKPERKEPFCGLKDSVNTSOKOELMKREKERNLDSRSLSRSISKERN 247  
 QY 220 VETLVVADDKMAAFHGA-GIKRYLLTVMAAARAFKRPISRNPSLVYTRVLVILGSGEEG 278  
 DB 248 VETLVVADVTKMIYHSENEVSYILTMNVGTGLFNPSIGNAHLVYVLLILLEEEG 307  
 QY 279 POGPSAOTLRFCAMORGLNTPEDSPDHFTALFTRODLCG--VSCDTLGMADVG 336  
 DB 308 LKIVHAERTLSFCWKQKINSKSDLPVHNDVAALLTRKDCAGFNRCETLGLSHLS 367  
 QY 337 TVCDPARSCAIVEDDGLSAFTAHNELGHVFNMLHDNS-KPCISLNGPLSTRHVAAPMAHV 395  
 DB 368 GKGQPHRSCINEDGLPLAFVAHGLHSGFIOHDSKENDG----EPGKRPFIMSPQLYD 423  
 QY 396 AHVDPEPSPCSARFTDPLDNGYGHCLDKPE 429  
 DB 424 QYDPTPLTWKSCSEYITRFLDRGWGLCLDDIPK 457

RESULT 14  
 AT10\_HUMAN STANDARD: PRT: 1077 AA.  
 AC 09H324:  
 DT 15-JUN-2002 (rel. 41, Created)  
 DT 15-JUN-2002 (rel. 41, Last sequence update)  
 DT 15-JUN-2002 (rel. 41, Last annotation update)  
 DE ADAMTS-10 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 10) (ADAM-TS 10) (ADAM-TS10) (Fragment).  
 GN ADAMTS10.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Apte S.S.;  
 RT "ADAM-TS10: a novel member of the ADAM-TS family containing multiple thrombospondin type 1 repeats."  
 RL Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.  
 CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (by similarity).  
 CC -1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE 1 DOMAINS ARE IMPORTANT FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX (By similarity).  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.  
 CC -1- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 5 TSP TYPE-1 DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: AF163762; AAC35563.1; -  
 CC MEROPS: M12.235; -  
 DR DR Genev; HGNC:13201; ADAMTS10.  
 DR DR InterPro: IPR001762; Disintegrin.  
 DR DR InterPro: IPR002870; Pep\_M12B\_propep.  
 DR DR InterPro: IPR001590; Reprolysin.  
 DR DR InterPro: IPR000884; TSP1.  
 DR DR InterPro: IPR000130; Zn\_M12peptidase.  
 DR DR Pfam: PF00090; tasp\_1; 5.  
 DR DR Pfam: PF01421; Reprolysin; 1.  
 DR DR Pfam: PF01562; Pep\_M12B\_propep; 1.  
 DR DR SMART: SMO0209; TSP1; 5.  
 DR DR PROSITE: PS02015; ADAM\_MPRO; 1.  
 DR DR PROSITE: PS50092; TSP1; 2.  
 DR DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 DR DR PROSITE: PS00427; DISINTEGRIN\_L; FALSE\_NEG.  
 KW Hydrolyase; Metalloproteinase; Zinc; Glycoprotein; Zymogen;  
 KW Repeat; Extracellular matrix.  
 FT FT NON\_TER 1  
 FT FT PROPEP 1  
 FT FT CHAIN <1 207 BY SIMILARITY.  
 FT FT METAL 208 1077 ADAMTS-10.  
 FT FT ACT\_SITE 366 366 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT FT METAL 370 367 BY SIMILARITY.  
 FT FT METAL 376 370 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT FT METAL 434 520 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT FT DOMAIN 478 520 DISINTEGRIN-LIKE.  
 FT FT DOMAIN 578 679 CYS-RICH.  
 FT FT DOMAIN 680 802 SPACER.  
 FT FT DOMAIN 521 577 TSP TYPE-1 1.  
 FT FT DOMAIN 799 860 TSP TYPE-1 2.  
 FT FT DOMAIN 862 918 TSP TYPE-1 3.  
 FT FT DOMAIN 922 976 TSP TYPE-1 4.  
 FT FT DOMAIN 961 1031 TSP TYPE-1 5.

FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 297 297 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 714 714 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 769 769 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 866 866 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SO SEQUENCE 1077 AA; 118072 MW; 3914DEL8DCBFB587 CRC64;  
 Query Match 19.9% Score 470.5; DB 1; Length 1077;  
 Best Local Similarity 31.9%; Pred. No. 8.7e-27;  
 Matches 136; Conservative 70; Mismatches 155; Indels 65; Gaps 16;  
 QY 60 EIVPEKLT--NGSVLPFGSGTP-----ARLCRLQAFGETLLLEEDSDGYOV 104  
 DB 14 ELAIPKRVHNGALLAFSPPPRRRGRTGATAESRLFKVASPTNHLNLTSSRLA 73  
 QY 105 EGLTVQYL--GQAPPELLGAEPTGLTGTINCDEPSVASLHMDGALLGVQYRAELH 161  
 DB 74 GHVSVYEWTRREGIAMQ--RAARPHCLYAGHLOGQASSHVAISTCGGLGLYVADEEYL 131  
 QY 162 LQPLEGTPNSAGP--GAHLIRKSPASGGP-----NCNK----- 197  
 DB 132 IEPLHGG-PKGRSPPEESPHVYVYKRS--SLRPHLDTACGYRDEKPKRGWMLFTLKP 188  
 QY 198 ---APLGSPSPRRRA-KRFASLSREFTLVYVADDKMAFHG-AGLRYLTVMAAARA 252  
 DB 189 PPARPLNGETRGOPOLKRSRERYETLVVADKKMVAHYGRDVEQYVLAIMNVAIL 248  
 QY 253 EKHPSIRNYSIVYRVLVILSGSEEGPQVPSAOTLRSCAMOR-----GLNTPED 304  
 DB 249 FQDSISGTVNLTLYRLLLTEDOPTLETHHAGSLDSFCWKMSIVNHSCHGNAPEN 308  
 QY 305 SPPDHDTFLITRODLGCVST--CDPLGMADVGVCDPARSCAIYEDDGLQSAFPAHE 362  
 DB 309 GVANH-DTAVILTRDICTLYKKPKCTGLAGVGGCEBERGCSVWEDGLQAFIAME 367  
 QY 363 LGHVENMLHNSKPCISLNGPLSTRHVNAVYMAHVPDEPSPSCARITFDLNGYGH 422  
 DB 368 ICHTFGMNHDGVGNSCGARGQ-DPAKLMAAHITMNTNP-FVWSSCNRDYITFLDSGLCL 425  
 QY 423 CLDKRP 428  
 DB 426 CLNRP 431  
 RESULT 15  
 AT06\_HUMAN STANDARD: PRT: 860 AA.  
 AC 090KPS:  
 DT 16-OCT-2001 (rel. 40, Created)  
 DT 16-OCT-2001 (rel. 40, Last sequence update)  
 DT 15-JUN-2002 (rel. 41, Last annotation update)  
 DE ADAMTS-6 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 6) (ADAM-TS 6) (ADAM-TS6).  
 GN ADAMTS6.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE-99395124; PubMed-10464288;  
 RA Hurekainen T.L., Hirohata S., Seldin M.F., Apte S.S.;  
 FT "ADAM-TS5, ADAM-TS6, and ADAM-TS7, novel members of a new family of Zinc Metalloproteinases."  
 RL J. Biol. Chem. 274:25555-25563(1999).  
 CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (by similarity).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED AT LOW LEVELS IN PLACENTA AND BARELY DETECTABLE IN A NUMBER OF OTHER TISSUES.  
 CC -1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.

CC -1- PIN: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY  
CC SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.  
CC -1- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.  
CC -1- SIMILARITY: CONTAINS 2 TSP TYPE-1 DOMAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: AF140674; AAD56357.1; -  
CC HSSP: P34179; 11NG.  
CC MEROPS: M12.230; -  
CC Genew: HGNC:222; ADAMT56.  
CC MIM: 605008; -  
CC InterPro: IPR001762; Disintegrin.  
CC InterPro: IPR002870; Pep\_M12B\_propep.  
CC InterPro: IPR001590; Reprolysin.  
CC InterPro: IPR000884; TSP1.  
CC InterPro: IPR000130; Zn\_Mtpeptidse.  
CC Pfam: PF00090; TSP\_1; 1.  
CC Pfam: PF01421; Reprolysin; 1.  
CC Pfam: PF01562; Pep\_M12B\_propep; 1.  
CC SMART: SM00209; TSP1; 1.  
CC PROSITE: PS50215; ADAM\_MEROP; 1.  
CC PROSITE: PS00142; ZINC\_PROTEASE; 1.  
CC PROSITE: PS50092; TSP1; 1.  
CC PROSITE: PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
CC Hydrolase: Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;  
KW Repeat: Extracellular matrix.  
ET SIGNAL 1 21 POTENTIAL.  
ET PROPEP 22 244 BY SIMILARITY.  
ET CHAIN 245 860 ADAMTS-6.  
ET METAL 403 403 ZINC (CATALYTIC) (BY SIMILARITY).  
ET ACT\_SITE 404 404 BY SIMILARITY.  
ET METAL 407 407 ZINC (CATALYTIC) (BY SIMILARITY).  
ET METAL 413 403 ZINC (CATALYTIC) (BY SIMILARITY).  
ET DOMAIN 453 509 ZINC (CATALYTIC) (BY SIMILARITY).  
ET DOMAIN 510 566 TSP TYPE-1 1.  
ET DOMAIN 567 668 TSP TYPE-1 1.  
ET DOMAIN 669 795 CYS-RICH.  
ET DOMAIN 796 852 SPACER.  
ET DOMAIN 852 852 TSP TYPE-1 2.  
ET DOMAIN 68 71 POLY-ARG.  
ET CARBOHYD 662 665 POLY-GLY.  
ET CARBOHYD 99 99 N-LINKED (GLCNAC. . .) (POTENTIAL).  
ET CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).  
ET CARBOHYD 222 222 N-LINKED (GLCNAC. . .) (POTENTIAL).  
ET CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).  
ET CARBOHYD 676 676 N-LINKED (GLCNAC. . .) (POTENTIAL).  
ET CARBOHYD 843 843 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 860 AA; 97098 MW; E57213015DECB2C5 CRC64;

Query Match 16.9%; Score 401; DB 1; Length 860;  
Best Local Similarity 29.6%; Pred No. 8,1e-22;  
Matches 121; Conservative 63; Mismatches 153; Indels 72; Gaps 12;

QY 34 LVMLLLLLLAST-LPSARLASPLPREETVFPPEKL-----NGSVLP----- 73  
DB 8 LVMILSLIMASSEFHDRLSYSSQEFVLYLEHYDLTPIPIVDQNGAFLSFTVKNDKHS 67  
QY 74 ----GSGTF-----ARLCRLQAFGETLLLEEDSDGVQVEGLTYOYLQ--APELLGCA 122  
DB 68 RRRRSMDPDPDPOAVSKLFLKSLAYGKHHLNLTNTDFVSKHFVLYEYWGDKGDPQKHDF 127  
QY 123 EPGTYLTGTINGPESVASLHMDGALLGLVQYRGAEHLHLPLEGSTPNSA-----GCP 176  
DB 128 LDNCHTGTGLQDRSTTKALSNCGVLRGVIAEDDEFIEPLKNTTEDSKHFSYENGHP 187  
QY 177 GAHLIRKRS-----PASGCGPMCNCVAKPLGSPSP-----RPRRAKRF 213

DB 188 --HVIYKKSALQORHLVDHSHCGVSDFTRSCKPMMIANDSTVSYSPLPNTNTHHROKRS 245  
QY 214 ASLSFVETLVVADDKMAAFHG--AGLKRYYLLTVMAAAKAFHPSIRPVSIVYTRVYL 272  
DB 246 VSIERFVETLVVADKMMVGYHGRKDIEHYLSVMNIVAKLYRDSLSLVNVIYARLVL 305  
QY 273 GSGEGAPQVGPSPAOTLRSFCAMQKGL-----NTPESDPDHFDFTALFTRODLGVS 325  
DB 306 TEDQPNLEINHADKSLDSFCWKQKSLTSHOSDQNTIPENGIAHHDNNAVLTIRYDICTYK 365  
QY 326 T--CDTLGMADVGVCDPARSCAIVEDDGLSFTAAHELGHVFMNLHD 372  
DB 366 NKPCGTLGLASVAGMCEPERSCSINEDIGLSAFTIAHEIYVHFGMNH 414

Search completed: April 21, 2003, 14:45:26  
Job time : 22.2332 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 14:43:36 ; Search time 18.5655 Seconds  
(without alignments)  
708.413 Million cell updates/sec

Title: US-10-050-200-8

Perfect score: 2370

Sequence: 1 MSQTGSHPRGRLAGRWLWGA.....PEAPLHLPVTGDKDDDDKG 447

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents-AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCITUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2313	97.6	837	US-09-122-126B-2	Sequence 2, Appl1
2	946	33.9	967	US-09-130-491-2	Sequence 2, Appl1
3	846.5	35.7	905	US-09-369-364A-9	Sequence 9, Appl1
4	661	27.9	930	US-09-122-126B-2	Sequence 2, Appl1
5	651.5	27.5	930	US-09-122-126B-15	Sequence 15, Appl1
6	639	27.0	245	US-09-369-364A-11	Sequence 11, Appl1
7	618	26.1	1882	US-09-130-491-13	Sequence 13, Appl1
8	600.5	25.3	608	US-09-130-491-13	Sequence 13, Appl1
9	559.5	22.3	874	US-09-369-364A-15	Sequence 15, Appl1
10	529.5	22.3	997	US-09-369-364A-7	Sequence 7, Appl1
11	401	16.9	859	US-09-369-364A-5	Sequence 5, Appl1
12	385	16.2	1224	US-09-930-872-4	Sequence 4, Appl1
13	364.5	15.4	566	US-09-491-522-7	Sequence 7, Appl1
14	361	15.2	491	US-09-930-872-2	Sequence 2, Appl1
15	358.5	15.1	1211	US-09-491-522-5	Sequence 5, Appl1
16	351.5	14.8	1081	US-09-369-364A-17	Sequence 17, Appl1
17	342.5	14.5	1205	US-09-491-522-11	Sequence 11, Appl1
18	222.5	9.4	439	US-09-026-001A-8	Sequence 8, Appl1
19	221.5	9.3	592	US-09-026-001A-14	Sequence 14, Appl1
20	218	9.2	521	US-09-026-001A-12	Sequence 12, Appl1
21	217	9.2	802	US-09-632-098-2	Sequence 2, Appl1
22	216.5	9.1	464	US-09-411-329C-14	Sequence 14, Appl1
23	208.5	8.8	462	US-09-411-329C-3	Sequence 3, Appl1
24	208.5	8.8	462	US-09-411-329C-17	Sequence 17, Appl1
25	206	8.7	621	US-09-026-001A-18	Sequence 18, Appl1
26	201	8.5	613	US-09-026-001A-10	Sequence 10, Appl1

28	200	8.4	621	US-09-026-001A-6	Sequence 6, Appl1
29	197.5	8.3	616	US-09-608-790-1	Sequence 1, Appl1
30	192.5	8.1	855	US-09-813-819-2	Sequence 2, Appl1
31	192.5	8.1	855	US-09-920-048-2	Sequence 2, Appl1
32	187	7.9	814	US-09-813-819-4	Sequence 4, Appl1
33	187	7.9	814	US-09-920-048-4	Sequence 4, Appl1
34	186.5	7.9	551	US-09-130-491-16	Sequence 16, Appl1
35	181.5	7.7	335	US-09-152-060-64	Sequence 64, Appl1
36	179	7.6	462	US-09-026-001A-16	Sequence 16, Appl1
37	178	7.5	529	US-08-836-442-3	Sequence 3, Appl1
38	177.5	7.5	769	US-08-243-542-4	Sequence 4, Appl1
39	177.5	7.5	769	US-08-477-407-4	Sequence 4, Appl1
40	177.5	7.5	769	US-08-484-355-4	Sequence 4, Appl1
41	169.5	7.2	751	US-08-836-443-3	Sequence 3, Appl1
42	168	7.1	391	US-08-706-216-6	Sequence 6, Appl1
43	168	7.1	470	US-08-813-150-2	Sequence 2, Appl1
44	162	6.8	201	US-09-411-329C-1	Sequence 1, Appl1
45	162	6.8	201	US-09-411-335-1	Sequence 1, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-122-126B-2  
; Sequence 2, Application US/09122126B  
; Patent No. 6451575  
GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: AGGREGATE DEGRADING METALLO PROTEASES  
; FILE REFERENCE: DM6909  
; CURRENT APPLICATION NUMBER: US/09/122,126B  
; CURRENT FILING DATE: 1998-07-24  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: Patencin version 3.0  
; SEQ ID NO. 2  
; LENGTH: 837  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-122-126B-2

Query Match 97.6%; Score 2313; DB 4; Length 837;  
Best Local Similarity 98.9%; Pred. No. 3.6e-219;  
Matches 439; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY	1	MSQTGSHPRGRLAGRWLWGAOPELLPIYPLSLVWLLLLLLASLPSRLASPLPREBE	60
DB	1	MSQTGSHPRGRLAGRWLWGAOPELLPIYPLSLVWLLLLLLASLPSRLASPLPREBE	60
QY	61	IYPERKLNQSVLPQSGTPARLLCRLQAFGETLLLEEDDSGVQVEGLTVQYGOAPELG	120
DB	61	IYPERKLNQSVLPQSGTPARLLCRLQAFGETLLLEEDDSGVQVEGLTVQYGOAPELG	120
QY	121	GAEPGYLYGTGTNGDESVASLHMDGALLGVLYQYGAELHQLPEGGTPNSAGGPGAH	180
DB	121	GAEPGYLYGTGTNGDESVASLHMDGALLGVLYQYGAELHQLPEGGTPNSAGGPGAH	180
QY	181	LRRKSPASQGPNCNKAKALGSPSPRRARRPASLSRVEVLTVVADDMKMAFHGGLKR	240
DB	181	LRRKSPASQGPNCNKAKALGSPSPRRARRPASLSRVEVLTVVADDMKMAFHGGLKR	240
QY	241	YLLTVAAAKAKFKHPSINPVSILVYTRVLIIGSGEGEQVGPSPAQTLRSCCAMOGLN	300
DB	241	YLLTVAAAKAKFKHPSINPVSILVYTRVLIIGSGEGEQVGPSPAQTLRSCCAMOGLN	300
QY	301	TPEDSDPDHFDYALLFTTRDLCGVSTCDTLGMAVDGTVCDPARSCAIVDDGLQSAFTAA	360
DB	301	TPEDSDPDHFDYALLFTTRDLCGVSTCDTLGMAVDGTVCDPARSCAIVDDGLQSAFTAA	360
QY	361	HELGHVFNMLHDSKRCISLNGPLSTSRHVMAVMAHVPEEPMSCSARFTTDFLDNXY	420
DB	361	HELGHVFNMLHDSKRCISLNGPLSTSRHVMAVMAHVPEEPMSCSARFTTDFLDNXY	420

QY 421 GHCLLDKPEAPLHLPVTDYKDD 444  
 DB 421 GHCLLDKPEAPLHLPVTDYKDD 444

# RESULT 2

US-09-130-491-2  
 ; Sequence 2, Application US/09130491  
 ; Patent No. 6416974  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Holtzman, Douglas A.  
 ; APPLICANT: Goedertl, Andrew D.J.  
 ; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83  
 ; FILE REFERENCE: 09404/041001  
 ; CURRENT APPLICATION NUMBER: US/09/130,491  
 ; CURRENT FILING DATE: 1998-08-07  
 ; EARLIER APPLICATION NUMBER: US 60/058,108  
 ; EARLIER FILING DATE: 1997-09-05  
 ; EARLIER APPLICATION NUMBER: US 60/054,961  
 ; EARLIER FILING DATE: 1997-08-06  
 ; NUMBER OF SEQ ID NOS: 16  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 2  
 ; LENGTH: 967  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-130-491-2

Query Match 39.9%; Score 946; DB 4; Length 967;  
 Best Local Similarity 45.7%; Pred. No. 2,8e-84;  
 Matches 208; Conservative 62; Mismatches 119; Indels 66; Gaps 14;

QY 37 LLLLLASLSPARLASPLPREEIVPEKINGSVLPGSTPARLCLQAFGETLLEL 96  
 DB 36 LLLAAALLAVSALORPBEDELYVPELERA---PGICT--RLHAFQOQDLEL 89  
 QY 97 EDDSGVQVEGLTVQYLGAPELIGAE---PGT---YLTGTINDPESVASLWDGGA 148  
 DB 90 RPDSSFLAGFTLQNYGRK---SGSETPLPETDLAHCFYSGTVGNDDPSAALLSCEG- 144  
 QY 149 LAGVLOYRGAELHLPLEGTPNSA--GPRGA-----HILRRK----- 184  
 DB 145 VRCAFLYLDIAVFTIOPLPASERLATAPEKRPAPLQFHLRRNRQGVGTGCVVDE 204  
 QY 185 -----SPASGQGPWCNVKAP---LGSPSPRRP-RAKRFASLSREYETLV 224  
 DB 205 PRPTGAETDEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDE 264  
 QY 225 VADDKAAAFHAGLAKRYLLTVMAAAKAFKHPISIRNPVSLVTVRLVILSGEERGPOVGPS 284  
 DB 265 VADQSMAEFHSGSLKHLYLLTFVSARLYKHPSIRNSVSLVYVKKILVHDEQKGEVTSN 324  
 QY 285 AAQTLRSFCAMORGCLNTPEDSDPHDTAILLFTRODLGVSCTDILGMADVGVCDPAPS 344  
 DB 325 AALLTANFCMVGKOHNPSPDRDREHYDTAILFTRODLGVSCTDILGMADVGVCDPAPS 384  
 QY 345 CAIVEDDGLOSAFTAHELGHVFMMLHDSKPCISLNGPLSTSRHYAVMAVMAVHDEEFP 404  
 DB 385 CSVIEDDGLOAFTAHELGHVFMNMHDAKQACASLNG--VNDOSHMAAMSLNDSQPM 443  
 QY 405 SPCSARFTDFDLNCGHCLLDKPEAPLHLPVTD 439  
 DB 444 SPCSAYMTISFLDNGHCLLDKPEAPLHLPVTD 476

## RESULT 3

US-09-369-364A-9  
 ; Sequence 9, Application US/09369364A  
 ; Patent No. 6391610  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Apte, Suneel  
 ; APPLICANT: Hurskainen, Tiina L.  
 ; APPLICANT: Hirahata, Satoshi

; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases  
 ; FILE REFERENCE: 26473/4007/10-30-00  
 ; CURRENT APPLICATION NUMBER: US/09/369,364A  
 ; CURRENT FILING DATE: 1999-08-06  
 ; NUMBER OF SEQ ID NOS: 31  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 9  
 ; LENGTH: 905  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus ADAMTS-8  
 ; US-09-369-364A-9

Query Match 35.7%; Score 846.5; DB 4; Length 905;  
 Best Local Similarity 41.2%; Pred. No. 1.6e-74;  
 Matches 192; Conservative 65; Mismatches 124; Indels 85; Gaps 15;

QY 36 WLLLLASLSPARLASPLP-----REELVPEKINGSVLPSPARLCLQAF 88  
 DB 10 WPLLLLLQLPPLPVCGAPAGPCTGAQASBLVPTRLPGSA-----SELAFLHLSAF 62  
 QY 89 GETLLELEDDSGVQVEGLTVQYLGAPELIGAE---GTUFGTINDPESVASLH-- 143  
 DB 63 GQGFVRLAPDMSFLAPERKIERLGSSNAAGS-EPGLNGCFSTGVNGERSLAMSCV 121  
 QY 144 --WDGCGALLGVLYRGAELHLPLEG---TPN--SAGCPGAILRRKSP----- 186  
 DB 122 AGWSSGFLLA---GEFTIOPGAGDSDLQPHRLQRMGPQ---RRDPGLAAAEVFP 173  
 QY 187 -----ASGQG-----PWCNVKAPLGSPSPRRAPKAF 213  
 DB 174 LPOGLEMEVEMGQGOQERSDNEEDKQDEKGLKETEDSRKVPFPFGSKT---RSKRF 229  
 QY 214 ASLSREYETLVVADDMKAFHAGLAKRYLLTVMAAAKAFKHPISIRNPVSLVTVRLVIG 273  
 DB 230 VSEARFVETLVADASMAAFGTGIDLNHLYVMAARLYKHPSIRNSVNLVYVKKLVE 289  
 QY 274 SGEERGPOVPSAOTLRSFCAMORGCLNTPEDSDPHDTAILLFTRODLGV--STCDTLGM 332  
 DB 290 KEMWGEVSDNGSLTLRNCNSQMRNRKPSDRHPEHYDTAILFTRODLGVSCTDILGM 349  
 QY 333 ADVGYCDPARSCAIVEDGLOSAFTAHELGHVFMMLHDSKPCISLNGPLSTSRHYVA 392  
 DB 350 ADVGTICDDPKSCSYIKKDGLOAAVYLAHELGHVLSMPHDDSKPCVRLGPMG-KYHMA 408  
 QY 393 PMAHVDPEEPPSPCSARFTDFDLNCGHCLLDKPEAPLHLPVTD 438  
 DB 409 PFPIHVNKTLPMSPCSAVYLTLLDDGHCILDAPTSVLP-IG 453

## RESULT 4

US-09-369-364A-2  
 ; Sequence 2, Application US/09369364A  
 ; Patent No. 6391610  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Apte, Suneel  
 ; APPLICANT: Hurskainen, Tiina L.  
 ; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases  
 ; FILE REFERENCE: 26473/4007/10-30-00  
 ; CURRENT APPLICATION NUMBER: US/09/369,364A  
 ; CURRENT FILING DATE: 1999-08-06  
 ; NUMBER OF SEQ ID NOS: 31  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 2  
 ; LENGTH: 930  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus ADAMTS-5  
 ; US-09-369-364A-2

Query Match 27.9%; Score 661; DB 4; Length 930;  
 Best Local Similarity 34.6%; Pred. No. 3.3e-56;  
 Matches 170; Conservative 69; Mismatches 160; Indels 92; Gaps 14;



```

; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 1882
; TYPE: PRF
; ORGANISM: Homo sapiens ADAMTS-9
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (468)
; OTHER INFORMATION: Xaa - C
; NAME/KEY: MOD_RES
; LOCATION: (521)
; OTHER INFORMATION: Xaa - Y
US-09-369-364A-13

Query Match      26.1%; Score 618; DB 4; Length 1882;
Best Local Similarity 35.2%; Pred. No. 1.7e-51;
Matches 169; Conservative 60; Mismatches 147; Indels 104; Gaps 18;

QY 31 LSWLVLLLLLLASLPASRLASP-----LPR-----EEIYVPEKLN-- 68
DB 6 VSWAT-LTLLVLDL--AEMGSPDAAAARVARDLHPROVKLETLESEYVSPIRVNL 61
QY 69 GSVLP-----GSGTPARLLCRLQAFGETLLELEQDSGV 102
DB 62 GEFPPNVHFKRTRRINSATDWPAPAFASSSSSTSSQAHYRLSATCOQFLPLTNAGF 121
QY 103 QVEGLTVQYLCAPPELLGCAEPCGTLYTGINDPESVASLHMDGALLGVLOYGAELHL 162
DB 122 IAPLFVTLTG-----TPGVNQTKEFS---EEEAELKHCYFRLCOYLRAHGRH- 168
QY 163 GPL---EGTPNSAGPGCAHLIRKSPASGQPMGVK---APLGS----- 202
DB 169 QPLRHEHNRKSKD-----KKTRARKWGRIRNLACDVALLNSGLATEAFSAVGNKT 221
QY 203 ---PSPRRPRAKRFASLSRFVETLVVADDKMAAFHAGLRYLLTVMAAAKAFKHPST 258
DB 222 DNTREKTRHRTKRFSLSYRFEVLVYADNRVSYGENLOHLLTLMSTVASTYDPSI 281
QY 259 RNPVSLVTRLVYLVSGEGEPGVGPSAOTLRSFCAMQKGLNTPEDSDPHEDTALIFTR 318
DB 282 GNLINIVNLVYIHNEODGPSISFNAQTLLKFKCOMHS--NSPGI---HHDVAVLLTR 337
QY 319 ODLCGV-STCDDTGMADVGVCOPARSCAIVEDDGLQSAFTAAHELGHVFNMLHDSKPC 377
DB 338 ODICRAHDKCDTLGLAEGLTICDPYRSCSISDSGLSTAFITAHELGHVFNMPHDNNKC 397
QY 378 ISLNGPLSTRHVMAPVMAHVDEEPWSPCSARFTFDLNGYGHCLLDKPEA-PLALPV 436
DB 398 --KEEVKSPQHMAPTLNFYTNPMWMSKSRKYTEFLDTGIGECLLNEPESRPPLPV 455

RESULT 8
US-09-130-491-13
; Sequence 13, Application US/09130491
; Patent No. 6416974
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US/09/130,491
; CURRENT FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: US 60/058,108
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: US 60/054,961
; EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13

```

```

; LENGTH: 608
; TYPE: PRF
; ORGANISM: Mus musculus
US-09-130-491-13

Query Match      25.3%; Score 600.5; DB 4; Length 608;
Best Local Similarity 61.4%; Pred. No. 1.6e-50;
Matches 105; Conservative 32; Mismatches 31; Indels 3; Gaps 2;

QY 269 LVTLGSGEGPGVGPASAOQLRSFCAMQKGLNTPEDSDPHEDTALIFTRDLCGVSTCD 328
DB 4 LIVTDEQKGEVTSNALTLRNFPCNMQKHPPSDORDAEHYDTALIFTRDLCGVSTCD 63
QY 329 TLGMADVGTCDPARSCAIVEDDGLQSAFTAAHELGHVFNMLHDSKPCISLNGPLSTR 388
DB 64 TLGMADVGTCDPSRCSVIEDDGLQSAFTAAHELGHVFNMPHDNAKOCASLNG-VNDS 122
QY 389 HWAPVMAHVDEEPWSPCSARFTFDLNGYGHCLLDKPEAPLPLVPTGD 439
DB 123 HMAWMSLMLDSHQSPWSPCSAYMTSFLDNGHGECLMDKRPQPIQLP--GD 171

RESULT 9
US-09-369-364A-15
; Sequence 15, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 874
; TYPE: PRF
; ORGANISM: Mus musculus ADAMTS-9
US-09-369-364A-15

Query Match      23.6%; Score 559.5; DB 4; Length 874;
Best Local Similarity 46.6%; Pred. No. 3.1e-46;
Matches 109; Conservative 41; Mismatches 73; Indels 11; Gaps 4;

QY 208 RRAKRPASLSRVEETLVVADDKMAAFHAGLRYLLTVMAAAKAFKHSINPVSIVYT 267
DB 122 KTRKRLSYPRVEVVVAVDHRVLYHGANLQHYITLMSIVASIKDSSIGNLINIVTY 181
QY 268 RLVYLGSGEGPGVGPASAOQLRSFCAMQKGLNTPEDSDPHEDTALIFTRDLC-GVST 326
DB 182 NLVYIHNEDEGPIYNNATTLKNTFCOMQHSKNY---LGIQHDTRVLYVTRDIDCAQDK 238
QY 327 CDTLGMADVGTCDPARSCAIVEDDGLQSAFTAAHELGHVFNMLHDSKPCISLNGPLST 386
DB 239 CDTLGLAEGLTICDPYRSCSISDSGLSTAFITAHELGHVFNMPHDNSKNC--KEGVKS 296
QY 387 SRHVAPVMAHVDEEPWSPCSARFTFDLNGYGHCLLDKPEA-----PLALPV 435
DB 297 PDIHMAPTLNFYTNPMWMSKSRKYTEFLDTGIGECLLNEPASRTYPLPSQLP 350

RESULT 10
US-09-369-364A-7
; Sequence 7, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00

```



```

RESULT 11
US-09-369-364A-5
; Sequence 5, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 859
; TYPE: prt
; ORGANISM: Homo sapiens ADAMTS-6
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (450)
; OTHER INFORMATION: Xaa = L
;
Query Match 16.9%; Score 401; DB 4; Length 859;
Best Local Similarity 29.8%; Pred. No. 1,3e-30;

```

```

Query Match 15.2%: Score 385; DB 4; Length 1224;
Best Local Similarity 25.8%: Pred. No. 8, 2e-29;
Matches 135; Conservative 68; Mismatches 170; Indels 150; Gaps

QY      35 VVLLLLLSTLLPSARL-----ASP-----LPR-----57
      1 : : : : : : : : : : : :
Db      14 LMLLLAOVAEQAPACAMGPAAALPSPSPVPPPPPPAERPPGMEKGEYDLSAYEVDRGD 73
      1 : : : : : : : : : : : :
QY      58 ---EEETVPEKANGSVLQSGCPARLLRLQAFGTTLLLEQSGVQVEGLTVQYGOA 119
      1 : : : : : : : : : : : :
Db      74 YSHLEIMHHQRRRAVAVSEVSLHL--RLKSGRHDPHVDLRTSSSLVAPRFIVYTLGKT 133
      1 : : : : : : : : : : : :
QY      116 PELLGAEPRGT-----YLTGTINDPESVASLHMDGALLGVLQYKCAEDLHLQ 165
      1 : : : : : : : : : : : :
Db      132 -----GTRSVQTLPPEDFCFGYQGLSRHSNSVALSTCOG--LSGMI RTEADYFLR 181
      1 : : : : : : : : : : : :
QY      164 PL-----EGGTPPSAGCGGAILRKS-----PASGO-----GPMC 194
      1 : : : : : : : : : : : :
Db      182 PLPSHLSMKTLGRNAGGSSP-----SHVILVKRSTEPHAPASAEVILVTSRTWELAHQPLH 234
      1 : : : : : : : : : : : :
QY      195 NKKAPLGSGS-----PPRRRAKRF-----ASLSRF-----VET 222

```



Db 263 DLEFDEYKSCLRHRSRLSRHNEELNVEVLVYDKKMONHENTTVYITLIMNV 322  
 QY 250 AKAFKPSIRNPVSLVTVLVLGSGEGRQYPSAQTLLRSPCAMORGINTPEDSDPH 309  
 Db 323 SALKFGTGGCNINLAIVGLILEDEQPGVLVSHHADHTLSSCCOMSGMGKDGTRHDH 382  
 QY 310 PFTALFTRODLGCVST--CDTLGMADVCTVCDPARSCAIVEDDGLQSAFTAHAHELGHVF 367  
 Db 383 ---ALLTGLDLCMKNECDTLGFAPISGMSKYSRSCINDETGLGIAFTLAHESGHNF 439  
 QY 368 NMLHD-NSKPCISLNGPLSTSRHWAPVMAHYDPEEPMSPCSAREFTYDPL 416  
 Db 440 GMHDEGNMCKKSE6-----NIMSPTLAGRNGVFSWSPCSROYLHKEL 483

## RESULT 15

US-09-491-522-5  
 ; Sequence 5, Application US/09491522  
 ; Patent No. 6428998  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Colige, Alain  
 ; APPLICANT: Lapiere, Charles M.  
 ; APPLICANT: Prockop, Darwin J.  
 ; TITLE OF INVENTION: RECOMBINANT N-PROTEINASE,  
 ; TITLE OF INVENTION: AND THE PRODUCTION, METHODS AND USES THEREOF  
 ; NUMBER OF SEQUENCES: 17  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pennie & Edmonds, LLP  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: NY  
 ; COUNTRY: USA  
 ; ZIP: 10036-2811  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: Windows  
 ; SOFTWARE: FastSeq for Windows Version 2.0b  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/491,522  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/886,333  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Abrams, Samuel B  
 ; REGISTRATION NUMBER: 30,605  
 ; REFERENCE/DOCKET NUMBER: 8389-0060-999  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 650-493-4935  
 ; TELEFAX: 650-493-5556  
 ; TELEX: 66141 PENNTE  
 ; INFORMATION FOR SEQ ID NO: 5:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1211 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; US-09-491-522-5

Query Match 15.1%; Score 358.5; DB 4; Length 1211;  
 Best Local Similarity 27.4%; Pred. No. 3.3e-26;  
 Matches 139; Conservative 63; Mismatches 183; Indels 123; Gaps 23;

QY 8 PGKGLAGRWLMGAPCLLPYPLSWLWLLLLLASLTP-----SARLAS-----P 54  
 Db 3 PPAGAAARLLC---PALLL-----LTLILPPILPPPPPANARLAAAADPPGCP 49  
 QY 55 LPREEF-----IVPEKLSVLPDGSCTPAR-----L 81  
 Db 50 LCHGASRIILAVPVRTDAQGLVSHVVSATSRAGVARRAAPVPTPSFFGNEEPGSHL 109

QY 82 LCRLOAFGETLLELEBODSGVQVEGLTVQVIGQ-----APELLGAEPGTLYLTGTING-- 134  
 Db 110 FYNVTYFAGSDILHLRLPNRLVAPGATMEMQEKGTTRVEPLLGS-----LYYGDVAGLA 165  
 QY 135 DPESVASLHMDGALLGLVQLYRGAELHLOPLECGTPNSAGCPG-AHILRRKSPASGQPM 193  
 Db 166 EASSVALSNCDG--LAGLIRMEEEFFIEPLEKGLAAQEAQGRVHVYRRPPTS----- 218  
 QY 194 CNVKAPLGSF-----SPPRRAKRRASLSRYE-TLVYA 226  
 Db 219 ---PPLGPGQALDTGASLDSLDSLSRALGVLEEHANSSRRARRHAADDVNEIVLLGV 274  
 QY 227 DDKMAAFHG-AGLRYLLTVMAAAAKAFKPSIRNPVSLVTVLVLGSGEGR--QVGP 283  
 Db 275 DDSVQFHGKEHQKYLTLTLMNIVNELYHDESGAHINIVVAILLSYKSMSLIEIG- 333  
 QY 284 SAAQTLLRSFCAMORGINTPEDSDPDHFDPAILFTRODLGCVSTCDTLGMADVCTVCDPAR 343  
 Db 334 NPSOSLENCVCRMAVYLQOKPDTGHDEYHDAIFLTRDPF-GPSGMQ--GYAPVTGMCHPVR 390  
 QY 344 SCAIVEDDGLQSAFTAHAHELGHVFNNLHD-NSKPCISLNGPLSTSRHWAPVMAHYDPEE 402  
 Db 391 SCTLNHEIDGSSAFVVAHETGHVGLGMEHGOQNRG---GDEVRLGSIAPLVOAAFHRE 446  
 QY 403 PMSPCSAREFTDPLDNGYGHCLDKPEA 430  
 Db 447 HNSRCSQOELSRYL-HSY-DCILIDPPA 472

Search completed: April 21, 2003, 14:49:13  
 Job time : 24.5655 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 14:29:41 : Search time 32.3706 Seconds  
(without alignments)  
1840.034 Million cell updates/sec

Title: US-10-050-200-8  
Perfect score: 2370  
Sequence: 1 MSQTSHPGRLAGRWLNGA.....PEAPLHPVTGDKDDDKG 447

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues  
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A.Geneseq\_101002:\*  
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2318	97.8	837	22	AAAG78228
2	2318	97.8	840	21	AAAB21256
3	2313	97.6	837	20	AAAW75425
4	2306	97.3	837	21	AAV99429
5	2306	97.3	837	22	AAU29199
6	2306	97.3	837	22	AAAB6178
7	1125	47.5	625	23	AAAB48394
8	979	41.3	896	21	AAAB21265
9	969.5	40.9	950	21	AAV53899
10	946	39.9	950	20	AAV49501

11	946	39.9	950	22	AAAB73549	Human ADAM-type me
12	946	39.9	950	22	AAAB50002	Human METH1. Homo
13	946	39.9	967	19	AAAB80285	Human integrin 11g
14	946	39.9	968	22	AAAB50011	Protein, SEQ ID 12
15	945	39.9	967	20	AAV04142	Human Tango-71 pro
16	936	39.5	967	20	AAW78189	Human secreted pro
17	898.5	37.9	950	22	AAAG62299	Human metalloprote
18	896.5	37.8	928	23	AAU72899	Human metalloprote
19	896.5	37.8	950	23	AAE22541	Human protease #2.
20	896.5	37.8	952	23	AAU74751	Human protease PR1
21	863.5	36.4	890	20	AAV49502	Human METH2. Homo
22	863.5	36.4	890	22	AAAB50003	Human ADAM type me
23	851	35.9	889	22	AAAB74946	Murine ADAMTS-8 am
24	846.5	35.7	905	22	AAAB72284	Rat metalloprotein
25	846	35.7	505	21	AAAB21257	Human ADAMTS-1 pro
26	798.5	33.7	727	20	AAW78435	Human metalloprote
27	703.5	29.7	680	21	AAAB21251	Human metalloprote
28	687	29.0	367	23	AAE22542	Human ADAMTS-5 am
29	661	27.9	930	22	AAAB72280	Murine ADAMTS-5 am
30	654	27.6	929	21	AAAB41226	Human ORFX ORF990
31	651.5	27.5	930	20	AAW75426	Human aggrecan deg
32	649.5	27.4	947	22	AAAB6950	Human metalloprote
33	649.5	27.4	1629	23	AAO14448	Human ADAMTS-SI pr
34	649.5	27.4	1935	23	AAU72896	Human metalloprote
35	648.5	27.4	870	21	AAAB21252	Rat metalloprotein
36	648.5	27.4	1629	23	AAAB30703	Human aggrecanase
37	643.5	27.2	1073	21	AAAB21264	Human metalloprote
38	642.5	27.1	1602	23	AAE30702	Human aggrecanase
39	641.5	27.1	1916	23	AAE19173	Human aggrecanase
40	639	27.0	245	22	AAAB72285	Human ADAMTS-8 am
41	635	26.8	1934	22	AAAB72301	Human ADAMTS-9 alt
42	633.5	26.7	381	21	AAAB21261	Human metalloprote
43	618	25.1	1882	22	AAAB72286	Human ADAMTS-9 am
44	606.5	25.6	1505	23	AAU72897	Human metalloprote
45	575.5	24.3	1907	23	AAU77133	Human metalloprote

## ALIGNMENTS

RESULT 1	
AAAG78228	
ID	AAAG78228 standard; Protein; 837 AA.
XX	
AC	AAAG78228;
XX	
DT	13-DEC-2001 (first entry)
XX	
DE	Human aggrecanase-1 SEQ ID NO 3.
XX	
KW	Human; aggrecanase-1; ADAMTS4; promoter; antiarthritic.
XX	
OS	Homo sapiens.
XX	
PN	JP2001245663-A.
XX	
PD	11-SEP-2001.
XX	
PF	06-MAR-2000; 2000JP-0059952.
XX	
PR	06-MAR-2000; 2000JP-0059952.
XX	
PA	(EISA ) EISA CO LTD.
XX	
DR	WPI; 2001-610072/70.
XX	
DR	N-PSDB; AA168146.
XX	
PT	New DNA for use as a promoter for controlling the expression of
XX	aggrecanase-1
XX	
PS	Disclosure; Page 9-11; 12pp; Japanese.
XX	
CC	The invention relates to a DNA acting as a promoter controlling the

CC expression of aggreganase-1 (ADAMTS4) resulting in antiarthritic  
CC activity. The DNA is used in a method for screening a compound affecting  
CC the expression control of the aggreganase-1 gene.

SO Sequence 837 AA:

Query Match 97.8% Score 2318; DB 22; Length 837;  
Best Local Similarity 99.1% Pred. No. 3.2e-198;  
Matches 440; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

OY 1 MSQTGSHPGGLAGRWLMGAOPCLLPYPLSWMLLLLLSLPSARLASPLPREEE 60
DB 1 MSQTGSHPGGLAGRWLMGAOPCLLPYPLSWMLLLLLSLPSARLASPLPREEE 60
OY 61 IYPEKINGSVLPGSGTPARLLCRLOAFGETLLELEODSGVVEGITYOYLGAPELLG 120
DB 61 IYPEKINGSVLPGSGTPARLLCRLOAFGETLLELEODSGVVEGITYOYLGAPELLG 120
OY 121 GAEPGYTLTGTINGDPESVASLHMDGALLGVLYRGAELHLQPLEGGTPNSAGPGCAHI 180
DB 121 GAEPGYTLTGTINGDPESVASLHMDGALLGVLYRGAELHLQPLEGGTPNSAGPGCAHI 180
OY 181 LRRKSPASGCGPMCNVAKAPLGSPSPRRRAKRFASLSRFVETLVADDKMAAFHGAGLKR 240
DB 181 LRRKSPASGCGPMCNVAKAPLGSPSPRRRAKRFASLSRFVETLVADDKMAAFHGAGLKR 240
OY 241 YLLTVMAAAAKAFKHPISIRNPVSLVYTRVLYLSGEGEPGVGSAQTLSFCAMRGSLN 300
DB 241 YLLTVMAAAAKAFKHPISIRNPVSLVYTRVLYLSGEGEPGVGSAQTLSFCAMRGSLN 300
OY 301 TPEDSDPDHFDTAILFTRODLGCVSTCDTLGMADVGVCDPARSCAIVEDDGLQSAFTAA 360
DB 301 TPEDSDPDHFDTAILFTRODLGCVSTCDTLGMADVGVCDPARSCAIVEDDGLQSAFTAA 360
OY 361 HELGHVFNMLHDSKPCISLNGPLSTRHYMAVMAHVDEEPMSCSARFITDFLDNGY 420
DB 361 HELGHVFNMLHDSKPCISLNGPLSTRHYMAVMAHVDEEPMSCSARFITDFLDNGY 420
OY 421 GHCLLDKPEAPLHLPTVGYDKDD 444
DB 421 GHCLLDKPEAPLHLPTVGYDKDD 444

```

RESULT 2

ID AAB21256 standard; Protein; 840 AA.

AC AAB21256;

DT 23-FEB-2001 (first entry)

DE Human metalloproteinase KIAA0688.

KW Human: KIAA0688; metalloproteinase; ADAM;  
KW a disintegrin and metalloproteinase domain; thrombospondin domain;  
KW vaccine; neurotrophic; neuroprotective; antiparkinsonian;  
KW ceretrolprotective; cytoskeletal; antiarthritic; immunosuppressive;  
KW Alzheimer's disease; Parkinson's disease; stroke; cancer; arthritis;  
KW autoimmune disease; brain tumour; brain injury.

OS Homo sapiens.

PN WO200053774-A2.

PD 14-SEP-2000.

PF 08-MAR-2000; 2000MO-US06237.

PR 08-MAR-1999; 99US-0264585.

PA (NEUR-) NEUROCRINE BIOSCIENCES INC.

PI Kelner GS, Clark M, Maki RA;

XX WPI: 2000-594326/56.  
DR N-PSDB: AAA05826.

XX

PT Polynucleotide encoding novel members of a disintegrin.

PS Claim 13; Fig 12; 129pp; English.

XX

CC The present sequence is human metalloproteinase KIAA0688. KIAA068  
CC was identified by as a member of the ADAMTS family of proteins by  
CC searching Genbank for sequences similar to ADAMTS-1 and ADAMTS-3. The  
CC ADAMTS family of proteins contains thrombospondin, integrin and  
CC metalloproteinase domains. ADAMTS polypeptides are useful for the  
CC manufacture of medicaments for treating conditions associated with  
CC neuroinflammation and/or neurodegeneration, such as Alzheimer's disease,  
CC Parkinson's disease and stroke. They are also useful for treating  
CC conditions associated with cell proliferation, cell migration,  
CC inflammation and/or angiogenesis, such as cancer, arthritis and  
CC autoimmune diseases. They can be used to treat patients afflicted with an  
CC invasive tumour, a brain tumour or brain injury.

SO Sequence 840 AA:

Query Match 97.8% Score 2318; DB 21; Length 840;  
Best Local Similarity 99.1% Pred. No. 3.2e-198;  
Matches 440; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

OY 1 MSQTGSHPGGLAGRWLMGAOPCLLPYPLSWMLLLLLSLPSARLASPLPREEE 60
DB 1 MSQTGSHPGGLAGRWLMGAOPCLLPYPLSWMLLLLLSLPSARLASPLPREEE 60
OY 61 IYPEKINGSVLPGSGTPARLLCRLOAFGETLLELEODSGVVEGITYOYLGAPELLG 120
DB 61 IYPEKINGSVLPGSGTPARLLCRLOAFGETLLELEODSGVVEGITYOYLGAPELLG 120
OY 121 GAEPGYTLTGTINGDPESVASLHMDGALLGVLYRGAELHLQPLEGGTPNSAGPGCAHI 180
DB 121 GAEPGYTLTGTINGDPESVASLHMDGALLGVLYRGAELHLQPLEGGTPNSAGPGCAHI 180
OY 181 LRRKSPASGCGPMCNVAKAPLGSPSPRRRAKRFASLSRFVETLVADDKMAAFHGAGLKR 240
DB 181 LRRKSPASGCGPMCNVAKAPLGSPSPRRRAKRFASLSRFVETLVADDKMAAFHGAGLKR 240
OY 241 YLLTVMAAAAKAFKHPISIRNPVSLVYTRVLYLSGEGEPGVGSAQTLSFCAMRGSLN 300
DB 241 YLLTVMAAAAKAFKHPISIRNPVSLVYTRVLYLSGEGEPGVGSAQTLSFCAMRGSLN 300
OY 301 TPEDSDPDHFDTAILFTRODLGCVSTCDTLGMADVGVCDPARSCAIVEDDGLQSAFTAA 360
DB 301 TPEDSDPDHFDTAILFTRODLGCVSTCDTLGMADVGVCDPARSCAIVEDDGLQSAFTAA 360
OY 361 HELGHVFNMLHDSKPCISLNGPLSTRHYMAVMAHVDEEPMSCSARFITDFLDNGY 420
DB 361 HELGHVFNMLHDSKPCISLNGPLSTRHYMAVMAHVDEEPMSCSARFITDFLDNGY 420
OY 421 GHCLLDKPEAPLHLPTVGYDKDD 444
DB 421 GHCLLDKPEAPLHLPTVGYDKDD 444

```

RESULT 3

ID AAW75425 standard; Protein; 837 AA.

AC AAW75425;

DT 30-MAR-1999 (first entry)

DE Human aggregan degrading metalloproteinase 1.

KW Human: aggregan degrading metalloproteinase; cartilage; proteoglycan;

KW interglobular domain; matrix metalloproteinase; bovine; interleukin-1beta;  
KW primer; PCR; amplification; inhibitor; cleavage; inhibitor; ADMP;  
KW arthritis; joint injury; pseudogout.  
XX  
XX Homo sapiens.  
XX  
XX WO9905291-A2.  
XX  
XX  
PD 04-FEB-1999.  
XX  
PF 24-JUL-1998; 98WO-US15438.  
XX  
XX 16-OCT-1997; 97US-0062169.  
PR 25-JUL-1997; 97US-0053850.  
PR 15-AUG-1997; 97US-0055836.  
XX  
XX (DDPO ) DU PONT PHARM CO.  
XX  
XX Arner EC, Burn TC, Copeland RA, Decicco CP, Liu R, Magolda R;  
PI Pratta M, Solomon KA, Tortorella MD, Tzaskos JM, Yang F;  
XX  
XX WPI: 1999-142943/12.  
DR N-PSDB: AAX00725.  
XX  
XX  
XX New isolated aggrecan degrading metallo proteinases - used to develop  
PT products for treating, e.g. osteoarthritis, joint injury, reactive  
PT arthritis, psoriatic arthritis or juvenile rheumatoid arthritis  
XX  
XX Claim 25: Page 61-62; 73pp: English.  
XX  
XX This sequence represents the human aggrecan degrading metalloproteinase 1  
CC (ADMP-1). ADMP-1 and ADMP-2 (AAW75426) are novel proteinases that cleave  
CC the aggrecan (a major cartilage proteoglycan) between residues  
CC Glu373-Ala374 of the interglobular domain (compared with cleavage between  
CC Asn341-Phe342 by the matrix metalloproteinases MMP-1, -2, -3, -7, -8, -9  
CC and -13). ADMP-1 and ADMP-2 were isolated and purified from the  
CC conditioned media of bovine nasal cartilage stimulated by  
CC interleukin-1beta. The purified proteins were partially sequenced and  
CC primers were synthesised based on the resultant amino acid sequences  
CC (AAX00727-X00732 for ADMP-1 and AAX00733-X00736 for ADMP-2). These were  
CC used to isolate the corresponding genes from cDNA. The ADMP polypeptides  
CC can be used for identifying inhibitors of ADMP activity which would  
CC prevent cleavage of the aggrecan core protein, thereby decreasing the  
CC loss of aggrecan from cartilage. Such inhibitors can be used for treating  
CC diseases such as osteoarthritis, joint injury, reactive arthritis, acute  
CC pyrophosphate arthritis (pseudogout), psoriatic arthritis and juvenile  
CC rheumatoid arthritis.  
XX  
XX  
XX Sequence 837 AA:  
SO  
Query Match 97.6%; Score 2113; DB 20: Length 837;  
Best Local Similarity 98.9%; Pred. No. 8.8e-198;  
Matches 439; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 301 TPEDSDPHDFRATLFTTRDCLGVCSTCPTLGMADYGTVCDDPARSCAIYEDDGLQSAFTAA 360  
DB 301 TPEDSDPHDFRATLFTTRDCLGVCSTCPTLGMADYGTVCDDPARSCAIYEDDGLQSAFTAA 360  
QY 361 HELGHVFNMLHDNSKPCISLNGPLSTRHVMAPVMAHVDPEEPMSPCSARFTTDFLDNGY 420  
DB 361 HELGHVFNMLHDNSKPCISLNGPLSTRHVMAPVMAHVDPEEPMSPCSARFTTDFLDNGY 420  
QY 421 GHCLLDKPEAPLPLPVTGDKDD 444  
DB 421 GHCLLDKPEAPLPLPVTGDKDD 444  
RESULT 4  
AAV99429  
ID . AAV99429 standard; Protein: 837 AA.  
XX  
XX AAV99429;  
XX  
XX 08-AUG-2000 (first entry)  
XX  
XX Human PRO1563 (UNQ769) amino acid sequence SEQ ID NO:317.  
DE  
XX  
XX Human; PRO polypeptide; membrane bound protein; receptor; diagnosis;  
KW transmembrane; secretion; immunoadhesion; pharmaceutical; screening.  
XX  
XX Homo sapiens.  
XX  
XX WO200012708-A2.  
XX  
XX 09-MAR-2000.  
XX  
XX 01-SEP-1999; 99WO-US20111.  
XX  
XX 01-SEP-1998; 98US-0098716.  
XX 01-SEP-1998; 98US-0098749.  
XX 01-SEP-1998; 98US-0098750.  
XX 02-SEP-1998; 98US-0098803.  
XX 02-SEP-1998; 98US-0098821.  
XX 02-SEP-1998; 98US-0098843.  
XX 02-SEP-1998; 98US-0098844.  
XX 09-SEP-1998; 98US-0099536.  
XX 09-SEP-1998; 98US-0099596.  
XX 09-SEP-1998; 98US-0099598.  
XX 09-SEP-1998; 98US-0099602.  
XX 09-SEP-1998; 98US-0099642.  
XX 10-SEP-1998; 98US-0099741.  
XX 10-SEP-1998; 98US-0099754.  
XX 10-SEP-1998; 98US-0099763.  
XX 10-SEP-1998; 98US-0099792.  
XX 10-SEP-1998; 98US-0099808.  
XX 10-SEP-1998; 98US-0099812.  
XX 10-SEP-1998; 98US-0099815.  
XX 10-SEP-1998; 98US-0099816.  
XX 15-SEP-1998; 98US-0100385.  
XX 15-SEP-1998; 98US-0100388.  
XX 15-SEP-1998; 98US-0100390.  
XX 16-SEP-1998; 98US-0100584.  
XX 16-SEP-1998; 98US-0100627.  
XX 16-SEP-1998; 98US-0100661.  
XX 16-SEP-1998; 98US-0100662.  
XX 16-SEP-1998; 98US-0100683.  
XX 17-SEP-1998; 98US-0100684.  
XX 17-SEP-1998; 98US-0100684.  
XX 17-SEP-1998; 98US-0100710.  
XX 17-SEP-1998; 98US-0100711.  
XX 17-SEP-1998; 98US-0100919.  
XX 17-SEP-1998; 98US-0100930.  
XX 18-SEP-1998; 98US-0100848.  
XX 18-SEP-1998; 98US-0100849.  
XX 18-SEP-1998; 98US-0101014.  
XX 18-SEP-1998; 98US-0101068.  
XX 18-SEP-1998; 98US-0101071.





```
RESULT 5
AAU29199
ID AAU29199 standard: Protein; 837 AA.
AC AAU29199;
XX
XX 18-DEC-2001 (first entry)
XX
XX Human PRO polypeptide sequence #176.
DE
XX PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
XX dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
XX blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
XX adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.
OS
XX Homo sapiens.
XX
XX WO20016848-A2.
XX
XX 20-SEP-2001.
XX
XX 28-FEB-2001; 2001WO-US06520.
XX
XX 01-MAR-2000; 2000WO-US05601.
XX 02-MAR-2000; 2000WO-US05841.
XX 03-MAR-2000; 2000US-187202P.
XX 06-MAR-2000; 2000US-186968P.
XX 14-MAR-2000; 2000US-189320P.
XX 14-MAR-2000; 2000US-189328P.
XX 15-MAR-2000; 2000WO-US06884.
XX 21-MAR-2000; 2000US-190828P.
XX 21-MAR-2000; 2000US-191007P.
XX 21-MAR-2000; 2000US-191048P.
XX 21-MAR-2000; 2000US-191314P.
XX 28-MAR-2000; 2000US-192655P.
XX 29-MAR-2000; 2000US-193032P.
XX 29-MAR-2000; 2000US-193053P.
XX 30-MAR-2000; 2000WO-US08439.
XX 04-APR-2000; 2000US-194449P.
XX 04-APR-2000; 2000US-194647P.
XX 11-APR-2000; 2000US-195875P.
XX 11-APR-2000; 2000US-196000P.
XX 11-APR-2000; 2000US-196187P.
XX 11-APR-2000; 2000US-196690P.
XX 11-APR-2000; 2000US-196820P.
XX 18-APR-2000; 2000US-198121P.
XX 18-APR-2000; 2000US-198585P.
XX 25-APR-2000; 2000US-199397P.
XX 25-APR-2000; 2000US-199550P.
XX 03-MAY-2000; 2000US-199654P.
XX 17-MAY-2000; 2000WO-US13705.
XX 22-MAY-2000; 2000WO-US14042.
XX 30-MAY-2000; 2000WO-US14941.
XX 02-JUN-2000; 2000WO-US15264.
XX 05-JUN-2000; 2000US-209832P.
XX 28-JUL-2000; 2000WO-US20710.
XX 22-AUG-2000; 2000US-0644848.
XX 24-AUG-2000; 2000WO-US23328.
XX 08-NOV-2000; 2000WO-US30952.
XX 01-DEC-2000; 2000WO-US32678.
XX 20-DEC-2000; 2000WO-US34956.
XX
XX (GENTH ) GENEINTECH INC.
XX
XX Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
XX Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX
XX WPI; 2001-602746/68.
XX
XX N-PSDB; AAS46100.
XX
XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the
```

```
PT presence of tumours, such as prostate and breast tumours, in mammals and
PT to screen for modulators of the compounds -
XX
XX Claim 11; Fig 352; 774pp; English.
XX
XX Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.
CC The PRO polypeptides and their associated nucleic acids can be used to
CC detect the presence of a tumour in a mammal by comparing the level of
CC expression of a PRO polypeptide in a test sample of cells from the animal
CC and a control sample of normal cells, whereby a higher level of
CC expression in the test sample indicates the presence of a tumour in the
CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats
CC and rabbits but are preferably human. The polypeptides can be used to
CC stimulate tumour necrosis factor (TNF) alpha release from human blood,
CC when contacted with it. A specific polypeptide can be used to stimulate
CC the proliferation or differentiation of chondrocyte cells. The PRO
CC proteins can be used to determine the presence of tumours and also
CC susceptibility to tumour development, particularly adrenal, lung, colon,
CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
CC can be used for genetic analysis of individuals with genetic disorders.
XX
XX Sequence 837 AA;
```

```
Query Match 97.3%; Score 2306; DB 22; Length 837;
Best Local Similarity 98.6%; Pred. No. 3.7e-197;
Matches 438; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1 MSQTGSHPRGLAGRWLGAQPCLLPIVPLSLWLLLLLLASLLPSARLASPLPREEE 60
DB 1 MSQTGSHPRGLAGRWLGAQPCLLPIVPLSLWLLLLLLASLLPSARLASPLPREEE 60
QY 61 IYPEEKLNSVLPSCGTPARLLCRLQAFETLLLEEDQSGVQVEBTLVQYLGQAPELLG 120
DB 61 IYPEEKLNSVLPSCGTPARLLCRLQAFETLLLEEDQSGVQVEBTLVQYLGQAPELLG 120
QY 121 GAEPCTYLTGTINGDESVASLHMDGALLGVLOYRGALHLQPLEGGTPNSAGGCAHI 180
DB 121 GAEPCTYLTGTINGDESVASLHMDGALLGVLOYRGALHLQPLEGGTPNSAGGCAHI 180
QY 121 GAEPCTYLTGTINGDESVASLHMDGALLGVLOYRGALHLQPLEGGTPNSAGGCAHI 180
DB 121 GAEPCTYLTGTINGDESVASLHMDGALLGVLOYRGALHLQPLEGGTPNSAGGCAHI 180
QY 181 LRRKSPASGQGPCNVKAPLGSPPRRRAKRPASLSRFEVLVVAADKMAAFHGAGLKR 240
DB 181 LRRKSPASGQGPCNVKAPLGSPPRRRAKRPASLSRFEVLVVAADKMAAFHGAGLKR 240
QY 241 YLLTYMAAAAKAFKHPSTINPVSIVYTRIVYLGSGEGQVQVPSAAQTLRSCAMRGIN 300
DB 241 YLLTYMAAAAKAFKHPSTINPVSIVYTRIVYLGSGEGQVQVPSAAQTLRSCAMRGIN 300
QY 301 TPEDSDPDHFDTAILFTRODLGCVSTCDTLGMADVTCVCDPARSCAIVEDDGLQSAFTAA 360
DB 301 TPEDSDPDHFDTAILFTRODLGCVSTCDTLGMADVTCVCDPARSCAIVEDDGLQSAFTAA 360
QY 361 HELGHVFNMLHDNSKPCISLNGPLSTSRHVMAPVMAHVDPEEPMSPCSARFTIDFLNCY 420
DB 361 HELGHVFNMLHDNSKPCISLNGPLSTSRHVMAPVMAHVDPEEPMSPCSARFTIDFLNCY 420
QY 421 GHCLLDKPEAPLHLVYTGDXKDD 444
DB 421 GHCLLDKPEAPLHLVYTGDXKDD 444
```

```
RESULT 6
AAB66178
ID AAB66178 standard: Protein; 837 AA.
XX
XX AAB66178;
XX
XX 02-APR-2001 (first entry)
XX
XX Protein of the invention #90.
XX
XX Secreted; transmembrane; gene therapy.
XX
```

OS Unidentified.  
 XX WO200078961-A1.  
 PN  
 XX 28-DEC-2000.  
 PD  
 XX  
 PF 18-FEB-2000; 2000WO-US04342.  
 XX  
 PR 23-JUN-1999; 99US-0141037.  
 PR 20-JUL-1999; 99US-0144758.  
 PR 26-JUL-1999; 99US-0145698.  
 PR 01-SEP-1999; 99MO-US20111.  
 PR 29-OCT-1999; 99US-0162506.  
 PR 30-NOV-1999; 99MO-US28313.  
 PR 02-DEC-1999; 99MO-US28551.  
 PR 16-DEC-1999; 99MO-US30095.  
 PR 05-JAN-2000; 2000WO-US00219.  
 PR 06-JAN-2000; 2000WO-US00376.  
 XX  
 XX (GENTH ) GENENTECH INC.  
 PA  
 XX Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S,  
 PI Gao W, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ,  
 PI Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D,  
 PI Matanabe CK, Williams PM, Wood WI;  
 XX WPI: 2001-071395/08.  
 DR  
 XX  
 XX Secreted and transmembrane proteins and nucleic acids designated PRO,  
 PT useful as hybridization probes; in chromosome and gene mapping and gene  
 therapy -  
 XX  
 XX Claim 1; Flg 180; 787bp; English.  
 XX  
 CC The present invention relates to secreted and transmembrane proteins.  
 CC These proteins and the DNA encoding them may be used as hybridization  
 CC probes in chromosome and gene mapping and in the generation of  
 CC anti-sense RNA and DNA. They may also be used to generate either  
 CC transgenic animals or knockout animals which are in turn useful for  
 CC development and screening of therapeutically useful reagents.  
 CC The nucleic acids may also be used in gene therapy.  
 CC  
 XX  
 XX  
 SQ Sequence 837 AA:  
 Query Match 97.3%; Score 2306; DB 22; Length 837;  
 Best Local Similarity 98.6%; Pred. No. 3.7e-197;  
 Matches 438; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 MSOTGSHPGRLAGRLMGAPCLLPILYPLSMIVLWLLLLLLASLPSRLASPLPREHE 60  
 DB 1 MSOTGSHPGRLAGRLMGAPCLLPILYPLSMIVLWLLLLLLASLPSRLASPLPREHE 60  
 QY 61 IYFPEKLNSVLPGSGTPARLLCRLOAFGETLLLEEDDSGVQVEGLTVQYGOAPELIG 120  
 DB 61 IYFPEKLNSVLPGSGTPARLLCRLOAFGETLLLEEDDSGVQVEGLTVQYGOAPELIG 120  
 QY 121 GAEPCTYLTGTINGDPESVSLHMDGALLGVLYRGAEHLIQLPLEGGTPNSAGCGAHI 180  
 DB 121 GAEPCTYLTGTINGDPESVSLHMDGALLGVLYRGAEHLIQLPLEGGTPNSAGCGAHI 180  
 QY 181 LRRKSPASGSGMNCVAKLGSPPRRARAFASLSRVEVLVYVADDMAFHGAGLKR 240  
 DB 181 LRRKSPASGSGMNCVAKLGSPPRRARAFASLSRVEVLVYVADDMAFHGAGLKR 240  
 QY 241 YLLTYWAAAKAFKHPISIRNPVSLVTRVLVLSGSEEGQVGPSPAQTLRSFCAMORGLN 300  
 DB 241 YLLTYWAAAKAFKHPISIRNPVSLVTRVLVLSGSEEGQVGPSPAQTLRSFCAMORGLN 300  
 QY 301 TPEDSDPDHFTAILFTRODLGCVSTCDTLGNADVGVCDPARSCAIVEDDGLQSAFTAA 360  
 DB 301 TPEDSDPDHFTAILFTRODLGCVSTCDTLGNADVGVCDPARSCAIVEDDGLQSAFTAA 360  
 QY 361 HELGIVFNNLHDSKPCISLNGPLSTSRHVMAVMAHVDPPEPMSPCSAFRTTDFLDNCGY 420

DB 361 HELGIVFNNLHDSKPCISLNGPLSTSRHVMAVMAHVDPPEPMSPCSAFRTTDFLDNCGY 420  
 QY 421 GHCLLDKPEAPLHLPTVTDGDKDD 444  
 DB 421 GHCLLDKPEAPLHLPTVTDGDKDD 444  
 RESULT 7  
 AAM48394  
 ID AAM48394 standard; Protein; 625 AA.  
 XX  
 AC AAM48394;  
 XX  
 DT 09-MAY-2002 (first entry)  
 XX  
 DE Rat aggreganase.  
 XX  
 KM Aggreganase; nontropic; neuroprotective; rat; Alzheimer's disease.  
 XX  
 OS Rattus sp.  
 XX  
 PN JP2001352991-A.  
 PD 25-DEC-2001.  
 XX  
 PF 16-JUN-2000; 2000JP-0181599.  
 XX  
 PR 16-JUN-2000; 2000JP-0181599.  
 XX  
 PA (DAUC ) DAICHI PHARM CO LTD.  
 XX  
 DR WPI: 2002-191859/25.  
 DR N-PSDB; ABA98676.  
 XX  
 XX A new gene, exhibiting aggreganase activity, induced by a beta-amyloid  
 PT protein, is useful in the preparation of drugs to treat Alzheimer's  
 PT disease.  
 XX  
 PS Claim 1; Page 12-13; 18pp; Japanese.  
 XX  
 CC The present sequence is a protein with aggreganase activity. This  
 CC sequence is rat-derived. The coding sequence for this protein can be used  
 CC in the preparation of drugs to treat Alzheimer's disease.  
 CC  
 XX  
 XX  
 SQ Sequence 625 AA:  
 Query Match 47.5%; Score 1125; DB 23; Length 625;  
 Best Local Similarity 91.8%; Pred. No. 1.1e-91;  
 Matches 213; Conservative 5; Mismatches 14; Indels 0; Gaps 0;  
 QY 213 FASLSRFVETLVVADDMKMAAFHAGLRYLLTVMAAAAKAFKHPISIRNPVSLVTRVLV 272  
 DB 1 FASLSRFVETLVVADDMKMAAFHAGLRYLLTVMAAAAKAFKHPISIRNPVSLVTRVLV 272  
 QY 273 GSGEGLPQVGPSPAQTLRSFCAMORGLNTPEDSDPDHFTAILFTRODLGCVSTCDTLGM 332  
 DB 61 GSGEGLPQVGPSPAQTLRSFCAMORGLNTPEDSDPDHFTAILFTRODLGCVSTCDTLGM 332  
 QY 333 ADVGVCDPARSCAIVEDDGLQSAFTAAHELGHVFNKMLHDSKPCISLNGPLSTSRHVMA 392  
 DB 121 AGGVGVCDPARSCAIVEDDGLQSAFTAAHELGHVFNKMLHDSKPCISLNGPLSTSRHVMA 392  
 QY 393 PYMAHVDPPEPMSPCSAFRTTDFLDNCGYCHCLLDKPEAPLHLPTVTDGDKDD 444  
 DB 181 PYMAHVDPPEPMSPCSAFRTTDFLDNCGYCHCLLDKPEAPLHLPTVTDGDKDD 444  
 RESULT 8  
 AAB21265  
 ID AAB21265 standard; Protein; 896 AA.  
 XX  
 AC AAB21265;

```

XX 23-FEB-2001 (first entry)
XX
XX Mouse metalloproteinase ADAMTS-1.
DE
XX Mouse; ADAMTS-1; metalloproteinase; ADAM;
XX a disintegrin and metalloproteinase domain; thrombospondin domain;
XX vaccinating; neurotrophic; neuroprotective; antiparkinsonian;
XX cerebrotective; cytosolic; antiarthritic; immunosuppressive;
XX Alzheimer's disease; Parkinson's disease; stroke; cancer; arthritis;
XX autoimmune disease; brain tumour; brain injury.
XX
XX Mus musculus.
OS
XX WO200053774-A2.
XX
XX 14-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US06237.
XX
XX 08-MAR-1999; 99US-0264585.
XX
XX (NEUR-) NEUROCRINE BIOSCIENCES INC.
XX
XX Keiner GS, Clark M, Maki RA;
XX
XX WPI; 2000-594326/56.
XX
XX Polynucleotide encoding novel members of a disintegrin,
XX metalloproteinase and thrombospondin domain protein family used to
XX prevent and treat Alzheimer's disease, cancer and autoimmune diseases -
XX
XX Disclosure; Fig 17; 129pp; English.
XX
XX The present sequence is mouse metalloproteinase ADAMTS-1. The
XX ADAMTS family of proteins is closely related to the ADAM (A Disintegrin
XX and Metalloproteinase Domain) family. Members of the ADAMTS family
XX contain a thrombospondin domain in addition to the disintegrin and
XX metalloproteinase domains found in the ADAMs. ADAMTS polypeptides are
XX useful for the manufacture of medicaments for treating conditions
XX associated with neuroinflammation and/or neurodegeneration, such as
XX Alzheimer's disease, Parkinson's disease and stroke. They are also
XX useful for treating conditions associated with cell proliferation, cell
XX migration, inflammation and/or angiogenesis, such as cancer, arthritis
XX and autoimmune diseases. They can be used to treat patients afflicted
XX with an invasive tumour, a brain tumour or brain injury.
XX
XX Sequence 896 AA:
SQ
Query Match 41.3%; Score 979; DB 21; Length 896;
Best Local Similarity 46.8%; Pred. No. 2,1e-78;
Matches 212; Conservative 66; Mismatches 109; Indels 66; Gaps 14;
OY 37 LLLLLL--LIPSAKLA--SPLPREBEIVPEKLVG--SCTPARLLCRLQAFGFT 91
DB 18 MLLLLLSTTMLCARGAGRTEDEELVPLSEKA--PGHDSITTRL--RLDAFGQO 72
OY 92 LLELEDSGVVEGLTVOYLGAPELLGGAE-----PGTYLGTCTINGDEPSVAS 141
DB 73 LHLKQLPPDSGFLAPGFTLQTVGRSP---GSAQHLDPDTGLAHCFYSCTVAGDPSASA 128
OY 142 LHMDSGALGLVLYRGAEHLQPLEG-----GTP--NSAGCGAHLIRKSPASGOG 191
DB 129 LSLGCG-VRGATYLDGEFFIOPARGVATERLAPVPEESSARQFHLRRRRGSG-G 186
OY 192 PKCNV-----KAPLGSSPPRR-----RAKRFASLSRYEET 222
DB 187 AKCGVMDDETLPTSDSPESQNTNRQWPYRDPDTPQDAGKPSGSGSIRKKRFVSSPRYET 246
OY 223 LVVADDKMAAFHAGIKRRLTVMAAAKAFKHPSTRNVSLVYRLVILSGGEGPQWG 282
DB 247 MLVAQDSMADFHSGKRLHLTLFVAAAFYKHPISIRNSISLVVKKILVIEYKQGPVET 306

```

```

OY 283 PSAOQLRSCFAMQGLNTPEDSDPDHFTATLFTRODLGVSCTGLMADVGTCPA 342
DB 307 SNAATLIRNFCMKOKHNSPSRDEPHYDTALLFTRODLGSHCTDTGLMADVGTCP 366
OY 343 RSCAIVEDDGLQSAFTAAHELGAVFNMJLHNSKPCISLNGPLSTRHWAPYMAHVDEE 402
DB 367 RSCSVIEDDGLQAAFTAAHELGAVFNMJLHNSKPCISLNGPLSTRHWAPYMAHVDEE 425
OY 403 PWSPCSARITFDLUNGCHCLDPEAPLHP 435
DB 426 PWSPCSAYWITSFLDNGHGECLMDKRPNDIKLP 458
RESULT 9
AAV53899
ID AAV53899 standard; Protein; 950 AA.
XX
XX AAV53899;
XX
XX 13-MAR-2000 (first entry)
XX
XX Amino acid sequence of a murine ADAMTS-1 protein.
XX
XX GON-1; metalloproteinase; cell migration; modulator; ADAMTS-1;
XX metalloproteinase domain; thrombospondin domain; abnormal cell migration;
XX organ shaping; sterility; cancer metastasis.
XX
XX Mus sp.
XX
XX WO9961656-A2.
XX
XX 02-DEC-1999.
XX
XX 28-MAY-1999; 99WO-US11918.
XX
XX 29-MAY-1998; 98US-0087170.
XX
XX 13-APR-1999; 99US-0129023.
XX
XX (WISC) WISCONSIN ALUMNI RES FOUND.
XX
XX Kimple JE, Blueloch RH;
XX
XX WPI; 2000-072633/06.
XX
XX Identifying modulators of proteins containing metalloproteinase and
XX thrombospondin domains, potentially useful for controlling cell
XX migration and organ shaping.
XX
XX Disclosure; Fig 1C; 60pp; English.
XX
XX The present sequence represents a murine ADAMTS-1 protein. ADAMTS-1 is
XX a metalloproteinase. The specification describes another related
XX metalloproteinase, a C. elegans GON-1 protein, that lacks a transmembrane
XX domain and possesses a predicted metalloproteinase domain between residues
XX 269-456. In C. elegans hermaphrodites, GON-1 is required for migration of
XX two distal tip cells to produce elongated tubes, whereas in males, GON-1
XX is required for migration of a single linker cell to produce a single
XX elongated tube. The protein is used in the method of the invention. The
XX specification describes a method for identifying a modulator of a
XX protein that contains a metalloproteinase domain and a thrombospondin
XX domain. The method comprises treating a target organism, having a
XX 'developing gonadal cell that is responsive to the protein, with a test
XX compound, and determining any change in migration or shape of the cell
XX attributable to the test compound. The compounds identified are
XX potential therapeutic modulators of abnormal cell migration and organ
XX shaping, e.g. for rendering animals (specifically nematodes) sterile
XX and for inhibiting cancer metastases.
XX
XX Sequence 950 AA:
SQ
Query Match 40.9%; Score 969.5; DB 21; Length 950;
Best Local Similarity 46.6%; Pred. No. 1.6e-77;
Matches 211; Conservative 68; Mismatches 107; Indels 67; Gaps 15;

```

-

Claim 10; Fig 1; 457bp; English.

Pt  
XX AA32000 and AA232001 encode, and AAY9501 and AAY9502 represent, human  
PS metalloproteinase thrombospondin (METH1) proteins METH1 and METH2  
CC respectively. METH1 and METH2 have been found to be potent inhibitors of  
CC angiogenesis both in vitro and in vivo. They can be used for treating  
CC cancer and other disorders related to angiogenesis including abnormal  
CC wound healing, inflammation, rheumatoid arthritis, psoriasis,  
CC endometrial bleeding disorders, diabetic retinopathy, some forms of  
CC macula degeneration, hemangiomas, and arterial-venous malformations.  
CC They may be useful in treating deficiencies or disorders of the immune  
CC system, by activating or inhibiting the proliferation, differentiation,  
CC or mobilisation (chemotaxis) of immune cells. The etiology of these  
CC immune deficiencies or disorders may be genetic, somatic, such as  
CC cancer or some autoimmune disorders, acquired (e.g., by chemotherapy or  
CC toxins), or infectious. They can also be used to treat inflammatory  
CC conditions, both chronic and acute conditions. The products can also be  
CC used for detection and diagnosis. AA232002 to AA232080, and AAY9503 to  
CC AY49511 represent sequences given in the exemplification of the present  
CC invention.

XX Sequence    950 AA;

Query Match                  39.9%; Score 946; DB 20; Length 950;  
Best Local Similarity         45.7%; Pred. No. 2,18-75;  
Matches    208; Conservative    62; Mismatches 119; Indels    66; Gaps    14;

OY      37 LLLLLLASSLSARLASLPREEELVPEPKLVGSCTPRALLCLQAFGELLLEL 96  
DB      19 LLLLAALAAVSADAGRSEEDDELVEPLERA--PGHG-TRL-RILHFDQDLDEL 72  
OY      97 EODSGVGEGVLIVGYLGCAPELLCGAE---PQT-----YLGTINGDPESVAASLIHDGA 148  
DB      73 RPDSSFLARGFTLOAVNGK----SSGETPIPETDIAHCFFYGSTGVGDSPSAALSICGG- 127  
OY      149 ILGLVQLYRGAEIHLIQPLEGCTPNNSA-GSGPA-----NILRK----- 184  
DB      128 VGCATYLLGEAVFIOTPLPASERLATAPGRKPAPRIOLPHLRNRQGDCVTGCVVDDE 187  
OY      185 -----SPASGOGPMKNVKAP-----IGSPPRP-RAKFPASISREVENTLY 224  
DB      188 PRPTSKAETEBEDEBECTBEDEBPONSPODDPALQGVGPRTGISTKKKFYSNHRIVEFML 247  
OY      225 VADDKMAAFHGAGLRYLLTYAAAATAKAFAKHRSIRNPVALTVTRLYTIGSGEGBOPWGS 284  
DB      248 VADQSMAEFHGSLGHXYLLTLFVSVARLYKHRSIKNSVSLVVYKLIVIHDOKGPREVISN 307  
OY      285 AAQITLRSCCANORGINTPEBDSDPHGFDAILLFRQDLGCVSTCDLTGMADVGTVCDDRARS 344  
DB      308 AALTTRLRNCSNMOKOHNPSPSRDRAEHYDTAILPTRRDDLCGSTCIDLGNADVGTCVCPERS 367  
OY      345 CAIVEDDCLQAFTTAHEHLGHVFNNLIHNNSKPCISLNCPLSTRSHVPANVMAHNDPEEPW 404  
DB      368 CSVIDEDDLQAAFTHAHELGHVFENNRRHODANOCASLNG-VNDSIMAMSMNLINDHSQPW 426  
OY      405 SPCSARFITDFLDNGYGHCLDKREPARYLHPVTGD 439  
DB      427 SPCSAYMITTSFLDNHGECLEMDKPPQNTPOLP--GD 459

RESULT\_11  
AAB73549  
ID    AAB73549 standard; Protein; 950 AA.  
XX  
XX    AAB73549;  
AC  
XT    07-AUG-2001    (first entry)  
DE  
KW Human ADAM-type metalloprotease MDT54, SEQ ID NO:4.  
Human; MDT54; ADAM-type metalloprotease; drug screening;

KW A Disintegrin And Metalloprotease; cancer; arthritis.  
 XX Homo sapiens.  
 OS  
 XX JP2001017183-A.  
 PN  
 XX 23-JAN-2001.  
 PD  
 XX 09-JUL-1999; 99JP-0196584.  
 PF  
 XX 09-JUL-1999; 99JP-0196584.  
 PR  
 XX (YAMA ) YAMANOUCHI PHARM CO LTD.  
 PA  
 XX WPI: 2001-275950/29.  
 DR N-PSDB; AAH20224.  
 XX  
 PT A new metal protease and its preparation for use as an anti-cancer and  
 XX anti-arthritis therapeutic -  
 XX  
 PS Claim 1; Page 12-14; 22pp; Japanese.  
 XX  
 CC The invention relates to the novel human ADAM (A Disintegrin And  
 CC Metalloprotease)-type metalloproteases MDTs4 (AAB73549) and MDTs5  
 CC (AAB73550). The metalloproteases can be used for the treatment of  
 CC cancers and arthritis. The invention also relates to the genes encoding  
 CC MDTs4 and MDTs5, vectors and host cells containing the MDTs4 or MDTs5  
 CC genes, the recombinant production of MDTs4 and MDTs5, and antibody  
 CC specific for MDTs4 or MDTs5, and methods of screening for compounds  
 CC which modulate the activity of MDTs4 and/or MDTs5. The present  
 CC sequence represents human MDTs4.  
 SO Sequence 950 AA;

Query Match 39.9%; Score 946; DB 22; Length 950;  
 Best Local Similarity 45.7%; Pred. No. 2.1e-75;  
 Matches 208; Conservative 62; Mismatches 119; Indels 66; Gaps 14;

37 LLLLLLASSLPASRLASPREEEIPEFKLNGSVLPAGSGTPARLRCRQAFGETLLEL 96  
 19 LLLLLAALLAVSALGRPSEDELEVPELERA---PGRGT-TRL-RLAHPQQLDEL 72  
 97 EDDSGVQVSGLTVOYLGAQPELLGAE---PCT-----YLTGTINDPESVASLHMDGGA 148  
 73 RPDSSFLACGFTLQNVGRK---SGSETPLPETDLAHCFYSGVNDGPDSSAALSLCEG- 127  
 149 LIGVLOVRCALHLQPLEGCTPNSA-CGPGA-----HILRRK----- 184  
 128 VRGAFYLLGEAYFIQPLPAASERLATAPGCKPPAPLQFILLRRNGDVGTCGVVDE 187  
 185 -----SPASGQGPCNVKAP---LQSPSPRP-RAKRFASLSRFVETIV 224  
 188 PRRTGAKEDEDEGTGEDEGQWSPDAPALOGVOPSTGSRKRRYSRRIVETML 247  
 225 VADDKAAAFHGAQKRYLLTVMAAAAKAKRHPRIKRPVSLVTVLILSGEGEPQVGP 284  
 248 VADQSAAEFGSGSLKHYLLTLFSAARLAKRHPISIRNSVSLVYVLIYHDEQGPETSN 307  
 285 AAOGLTSPCAMQGLTTPESDPDHDPTALFTRODLGCVSTDTGLMADVGYCDPARS 344  
 308 AALTTLNFCNMQRQHPSPDRDAEHDTALFTRODLGCGSQCTDITGLMADVGYCDPSRS 367  
 345 CAIVEDDGLQSAFTAAHGLGVNFMNLHDNSKPCISLNGPSTSRHVAAPVMAVDEPEW 404  
 368 CSVIEDDGLQAAATTAAHGLGVNFMNHDHDAKQASLNG-VNODSHMAASLHSLNDSPM 426  
 405 SPCSARFTDFDLNGYGHCLLDKPEAPLHLPTVGD 439  
 427 SPCSAYMITSFLDNGHGECLMDKRPNDIQP--GD 459

RESULT 12  
 AAB50002

ID AAB50002 standard; Protein; 950 AA.  
 XX  
 AC AAB50002;  
 XX  
 DT 19-MAR-2001 (first entry)  
 XX  
 DE Human METH1.  
 XX  
 KW Human; METH1: metalloprotease; thrombospondin; angiogenesis inhibition;  
 KW cancer therapy; benign tumour; ocular angiogenic disease;  
 KW rheumatoid arthritis; psoriasis; wound healing; endometriosis;  
 KW vasculogenesis; granululation; hypertrophic scar; nonunion fracture;  
 KW scleroderma; trachoma; vascular adhesion; myocardial angiogenesis;  
 KW coronary collateral; cerebral collateral; arteriovenous malformation;  
 KW ischaemic limb angiogenesis; Osler-Webber syndrome;  
 KW plaque neovascularisation; telangiectasia; haemophilic joint;  
 KW angiodiroma; fibromuscular dysplasia; wound granululation;  
 KW Crohn's disease; atherosclerosis; birth control.  
 KW  
 OS Homo sapiens.  
 XX  
 PN WO200071577-A1.  
 XX  
 PD 30-NOV-2000.  
 XX  
 PF 25-MAY-2000; 2000MO-US14462.  
 XX  
 PR 25-MAY-1999; 99US-0318208.  
 PR 20-JUL-1999; 99US-0144882.  
 PR 10-AUG-1999; 99US-0147823.  
 PR 13-AUG-1999; 99US-0373658.  
 PR 22-DEC-1999; 99US-0171503.  
 PR 22-FEB-2000; 2000US-0183792.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (SMR ) SMITHKLINE BECHAM CORP.  
 PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.  
 PA (IRUE/) IRUELA-ARISPE L.  
 PA (HAST/) HASTINGS G A.  
 PA (RUBE/) RUBEN S M.  
 PA (JONK/) JONAK Z L.  
 PA (TRUL/) TRULLI S H.  
 PA (FORN/) FORNWALD J A.  
 PA (TERR/) TERRETT J A.  
 XX  
 PI IrueLa-Arispe L, Hastings GA, Ruben SM, Jonak ZL, Trulli SH;  
 PI Fornwald JA, Terrett JA;  
 DR  
 DR N-PSDB; AAC90057.  
 XX  
 PT METH1 and METH2 polynucleotides and encoded polypeptides, used to  
 PT inhibit angiogenesis in the treatment of disorders such as cancer,  
 PT rheumatoid arthritis and psoriasis -  
 XX  
 PS Claim 15; Fig 1; 768pp; English.  
 XX  
 CC The present sequence is human METH1 (ME for metalloprotease and TH for  
 CC thrombospondin). METH1 can be used for inhibiting angiogenesis in an  
 CC individual, and for treating cancer, benign tumours, an ocular angiogenic  
 CC disease, rheumatoid arthritis, psoriasis, delayed wound healing,  
 CC endometriosis, vasculogenesis, granululations, hypertrophic scars, nonunion  
 CC fractures, scleroderma, trachoma, vascular adhesions, myocardial  
 CC angiogenesis, coronary collateral, cerebral collateral, arteriovenous  
 CC malformations, ischaemic limb angiogenesis, Osler-Webber syndrome, plaque  
 CC neovascularisation, telangiectasia, haemophilic joints, angiodiroma,  
 CC fibromuscular dysplasia, wound granululation, Crohn's disease or  
 CC atherosclerosis. METH1 can also be used in birth control. METH1 can also  
 CC be used in diagnostic methods for the prognosis of cancer.  
 XX  
 SO Sequence 950 AA;

Query Match 39.9%; Score 946; DB 22; Length 950;









2/	48	3.2	2848	4	US-09-309-304A-4	sequence 4, Appl
----	----	-----	------	---	------------------	------------------

36 / AGGCGGACGAGTGCCTCTGGCCCAACCGAGCCACTGCTTCTATCGGGGCACAGTGGG 426

Db	471	AGGGGGAGCGATGGCGCCCTTGGGCCACCGSAGACCACCTCTTCTATTCGGGGACAGTGA	530
Qy	427	CGGTATGTCGCCGCTCTTGCGCTGCTTTTGACCTCTGTGGGGGTCTCGACGGCTTCTTGC	486
Db	531	CGGTATGTCGCCGCTCTTGCGCTGCTTTTGACCTCTGTGGGGGTCTCGACGGCTTCTTGC	590
Qy	487	GGTCAACGACGGCGGCTACACCCCTAAAGCCACATGCTGGCGGAGCCCTGGCGGAGAGA	546
Db	591	GCTCAACGACCGCGGCTACACCCCTAAAGCCACATGCTGGCGGAGCCCTGGCGGAGAGA	650
Qy	547	AAAGGGCGCGGTGTACGGGGATGGTCCGACGGATCTGTACGCTTACACCCCGGAGGG	606
Db	651	AAAGGGCGCGGTGTACGGGGATGGTCCGACGGATCTGTACGCTTACACCCCGGAGGG	710
Qy	607	CTTGAGTTGGAGGCCCTTGCGCGCGGCGCGACGTGCGAAACCCCGGTCCACACGGGA	666
Db	711	CTTGAGTTGGAGGCCCTTGCGCGCGGCGCGACGTGCGAAACCCCGGTCCACACGGGA	770
Qy	667	GGCCCGACGACATGCTCCGGCGCAGACAAACCCGAGCGGACCGCAGCAGCTGGCTTCGA	726
Db	771	GGCCCGACGACATGCTCCGGCGCAGACAAACCCGAGCGGACCGCAGCAGCTGGCTTCGA	830
Qy	727	GCTCTTGAGACCATGTCGCCCTCTCTCGCCGCTGGGGGCTTAGACCGGACAGCTGGTGGC	786
Db	831	GCTCTTGAGACCATGTCGCCCTCTCTCGCCGCTGGGGGCTTAGACCGGACAGCTGGTGGC	890
Qy	787	GGGGGGCGCGGCTCATCTCCGGGCGCGCAGGTGAGGTGCTTGTGGTGGTGAAGC	846
Db	891	GGGGGGCGCGGCTCATCTCCGGGCGCGCAGGTGAGGTGCTTGTGGTGGTGAAGC	950
Qy	847	GTCATGGCGCGTTGTATGCCCCGGGGCCTGCAGCATTAACCTGTGACCTGGCCTCCAT	906
Db	951	GTCATGGCGCGTTGTATGCCCCGGGGCCTGCAGCATTAACCTGTGACCTGGCCTCCAT	1010
Qy	907	CGCCAAATAGGCTGTACAGCCATCTTAGCATATCGAAGAACATCCGCTGGCGGTGTGAA	966
Db	1011	CGCCAAATAGGCTGTACAGCCATCTTAGCATATCGAAGAACATCCGCTGGCGGTGTGAA	1070
Qy	967	GGTGGTGGTGTAGCGCAGAGCAAGAGCCTGTGAAGTGAAGCAAGAACGCTGCCACAC	1026
Db	1071	GGTGGTGGTGTGTAGCGCAGAGCAAGAGCCTGTGAAGTGAAGCAAGAACGCTGCCACAC	1130
Qy	1027	ACTCAAAACCTTTTGCAAGTGGAGAGCAACAACAACGAGCTGGAGATGACATGAGGA	1086
Db	1131	ACTCAAAACCTTTTGCAAGTGGAGAGCAACAACAACGAGCTGGAGATGACATGAGGA	1190
Qy	1087	GCACATGAGATGACAGCTATCCCTGTTTACTCGGGAGATTTATGTGGGCATCATTTCTGA	1146
Db	1191	GCACATGAGATGACAGCTATCCCTGTTTACTCGGGAGATTTATGTGGGCATCATTTCTGA	1250
Qy	1147	CACCGTGGGAATGGCAGACGTTGGGACCAATATGTTCTCAAGAGCCAGCTGTGCTGTAT	1206
Db	1251	CACCGTGGGAATGGCAGACGTTGGGACCAATATGTTCTCAAGAGCCAGCTGTGCTGTAT	1310
Qy	1207	TGAAGAGATGAGGCTTCCACGACCACTTCACTGTGGGTCAAGAAATCGACATTTACTTGG	1266
Db	1311	TGAAGAGATGAGGCTTCCACGACCACTTCACTGTGGGTCAAGAAATCGACATTTACTTGG	1370
Qy	1267	CTCTGCCATGAGGATTCCAAATTTCTGTGAAGAGACCTTTGCTTCACAGAAATAGCG	1326
Db	1371	CTCTGCCATGAGGATTCCAAATTTCTGTGAAGAGACCTTTGCTTCACAGAAATAGCG	1430
Qy	1327	CTTAAATGTCATCTCTTTCAGAGCAATGATGATCTAAGCCCTGGTCCAAATGACATTC	1386
Db	1431	CTTAAATGTCATCTCTTTCAGAGCAATGATGATCTAAGCCCTGGTCCAAATGACATTC	1490
Qy	1387	AGCCACCATCAGAAATTCCTGTGATGATGGCCATGTGTAACCTTTTGTCTGACCTTACACG	1446
Db	1491	AGCCACCATCAGAAATTCCTGTGATGATGGCCATGTGTAACCTTTTGTCTGACCTTACACG	1550
Qy	1447	AAAGCATCTCTGGCGGGGA 1467	

[illegible]



```

Db 1619 TCATAGAGATATGGTTTACAGCTGCTTCACACACACCATGATTAAGCCAGCTGT 1678
Qy 1263 TTGGCTCTCCATGACGATTCGCAAAATTCGTGAAGAGACCTTTGGTCCACAGAGATA 1322
Db 1679 TTAACTATGCCACATATGATGTGCAAAAGCAGTGTGCCACTTAATGTGTGAACAGAGATT 1738
Qy 1323 AGCGCTTAATGTCTTCATCTTACACAGCATGTATGATCAATGAGCCCTGGTCAATATGA 1382
Db 1739 CCACATGATGGCGTCAATATGCTTCCACACCTGACACAGCCAGCCCTGGTCCCTTGA 1798
Qy 1383 CTTACGCCACCATCAGAGAAATTCCTGATGATGGCATGTACTGTGTCGACCTTAC 1442
Db 1799 GTGCTACTACATGATTAATCATTTCTGTGATAATGTCTGGGAAATGTTTATGATGACAAC 1858
Qy 1443 CACGAAAGCAGATCTGT 1459
Db 1859 CTCAGATATCCCATACAG 1875

```

## RESULT 4

```

US-09-392-184-1
: Sequence 1, Application US/09392184
: Patent No. 635889
: GENERAL INFORMATION:
: APPLICANT: Robison, Keith E.
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN
: FILE REFERENCE: 5800-55
: CURRENT APPLICATION NUMBER: US/09/392,184
: CURRENT FILING DATE: 1999-09-09
: NUMBER OF SEQ ID NOS: 33
: SOFTWARE: FASTSEQ for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 4858
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(4858)
: OTHER INFORMATION: repolysin (ADAM family of metalloprotease)
: NAME/KEY: misc_feature
: LOCATION: (1)...(4858)
: OTHER INFORMATION: n - A,T,C or G
US-09-392-184-1

```

```

Query Match          16.9%  Score 256; DB 4; Length 4858;
Beat Local Similarity 53.4%; Pred No. 1.2e-46;
Matches 586; Conservative 0; Mismatches 505; Indels 7; Gaps 2;

```

```

Qy 363 CAGGAGCGGAGCAGATGCGCCCTGCGCCACCGAGCCACTGCTTCTATCGGGCAGAG 422
Db 779 CCGGGTCCGAGAGCGCGCTCCGGAACCGACCTGGCGACATGCTTCTATCGGGCAGCG 838
Qy 423 TGGACGCTAGTCCCGCTCTGCTGTCTTTGACCTCTGTGGGGGTCTCGACGGCTTCT 482
Db 839 TGAATGCGCATCTCAGCTGCTGCGCCCTGACCTCTGCGAGGGCGTGGCCGGCGCT 898
Qy 483 TCGCGGTCAAGCAGCGCGCTACACCTTAAGCCACTGCTGGCGGAGACCTTGGGGGAG 542
Db 899 TCTACCTGCTGGGGAGGGGATTTTATATCAGCGCTGCGCCCGCCAGCGAGCGCTCG 958
Qy 543 AAGAAAAGGGGCGCTGTATGCGGGATGGGTCCGACAGATCTCTGACGCTTACACCCGCG 602
Db 959 CCACGCGCGCCCGCAGGGAGAAAGCGCGCGGACACACTACAGTTTCAACCTCTGCGGGCGGA 1018
Qy 603 AGGCTTACGCTTCCAGAGCCCTGCGCGCGCGGCGACAGCTGCGAACCCTCGGCTCCACAC 662
Db 1019 ATCGGCGAGGGCGCAGTAGCGCGGACCTGCGGGGTGCTGAGACGACGAGCCCGCGGACTG 1078
Qy 663 GCGAGGCGCCACAGAGATGCTCGGGCGCACAGCAACCCGAGCGGAGCGGAGCAGTGGCT 722
Db 1079 GGAAGCGGAGACCGAGACGAGAGCGAAGGAGCTGAGGGCGAGGACGAAAGGCGCTCACT 1138

```

```

Qy 723 CGCAGCTCTTGGACAGATGCCGCTCTGCGCCGCTGGGGGCTCAGAGACCGCAGAGCTGT 782
Db 1139 GGTGCGCGGACAGACCGCGGACACTGCAAGGCGTGAAGACGACCCACAGAGA-----ACTGGAA 1192
Qy 783 GCGCGCGCGCGCGCGCTCTCATCTCCCGGGCCCGCAGGTGAGACCTGCTTGTGTGGCTG 842
Db 1193 GCATTAAGAAAGAAAGGATTTGTGTCCAGTCACCGGCTATGTGGAAACCATGCTTGTGGCGAG 1252
Qy 843 ACCGCGTCTGCGCGGCGGTGTATGGCGGGCGCTGACGATTAACCTGCT -GACCGTGGCC 901
Db 1253 ACCAGTCTGATGCGCAGATTTCCACGGCAGTGTCTTAAGCATTAACCTTCTCAACGTTGTTT 1312
Qy 902 TCCATGCGCCAAATAGGCTGTACAGCCATGCTTACATGAGAACACATCCGCTGGCCGTG 961
Db 1313 TCGGTGGCAGCCAGATTTGTACAAACACCCCGACATTTCTTAATTCAGTTAGTTAGTGGTGG 1372
Qy 962 GTGAAGGTGGTGTGCTAGGCGCAAGACAGAGAGCTGGAAAGTGAAGCAAGACCTGCG 1021
Db 1373 GTGAAGATCTTGTGTATCCAGCATGACAGAAAGGGGCGGAGAGTACCTTCAATGCTGCC 1432
Qy 1022 ACCACACTCAAGACTTTTGCAGTGGCGACGACCAACACACACAGCTGGAGATGACCAT 1081
Db 1433 CTCACCTCTGCGAATCTTTTGCACTGGCAGAGACACACACACACACGATACCGGGAT 1492
Qy 1082 GAGAGCACTACGATGACAGCTATCTGTTTACTCGGAGAGATTATGTGGCATCATTTCA 1141
Db 1493 GCAGAGCACTATGACACAGCAATTTCTTTTACACAGACAGCTGTGTGGTCCAGACA 1552
Qy 1142 TGTGACACCTCTGGGAATGGCAGACAGTGGGACCATATGTTCTCCAGAGCGCAGCTGTGT 1201
Db 1553 TGTGATACCTTGGAGATGCTGATGTGGAACTGTGTGATGCCAGAGAAAGTGTCTC 1612
Qy 1202 GTGATTTAAGAGAGATGGCTCCGACGAGCCTTCACTGTGGCTCAGCAAAATCGACATTTA 1261
Db 1613 GTCATTAAGAGATGATGTTTACAGCTGCTTCCACAGACCCCATGAATTAAGCCACGTG 1672
Qy 1262 CTTGGCTCTCCATGACAGATTCGCAAAATTCGTGAAGACACCTTTGGTTCACAGAAAT 1321
Db 1673 TTTAACATGCGCATGATGATGCAAAAGCAGTGTGCCAGCTTAATGTGTGAACAGAGAT 1732
Qy 1322 AAGCGCTTAATGTTCTTCCATCTTACAGCATTTGATGATCACTTAAGCCCTGGTCCAAATGC 1381
Db 1733 TCCACATGATGCGGTCAATGCTTTCCAACTGGACACAGCAGCAGCCTTGTCTCTTGC 1792
Qy 1382 ACTTACGCCACCATCAGCAAAATTCCTGATGATGGCATGTAACTGTTTGTGTCGACCTA 1441
Db 1793 AGTGCTTACATATATATCATATTTCTGATTAATGTGATGGGAATGTTTGAATGACAAAG 1852
Qy 1442 CCAGCAAGCAGATCTGT 1459
Db 1853 CCTCAGAAATCCCATACAG 1870

```

## RESULT 5

```

US-09-484-970B-58
: Sequence 58, Application US/09484970B
: Patent No. 6426186
: GENERAL INFORMATION:
: APPLICANT: Jones, Karen A.
: APPLICANT: Volkmuth, Wayne
: TITLE OF INVENTION: BONE REMODELING GENES
: FILE REFERENCE: PB-0014 US
: CURRENT APPLICATION NUMBER: US/09/484,970B
: CURRENT FILING DATE: 2000-01-18
: NUMBER OF SEQ ID NOS: 172
: SOFTWARE: PERL Program
: SEQ ID NO 58
: LENGTH: 3706
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature

```

```

: OTHER INFORMATION: Incyte ID NO. 6426186 007074.1
: NAME/KEY: unsure
: LOCATION: 3634, 3638-3639, 3642-3643, 3647-3648, 3652, 3654-3658, 3664, 3674, 3681
: OTHER INFORMATION: a, t, c, g, or other
US-09-484-970B-58

```

```

Query Match          16.6%  Score 251.8; DB 4; Length 3706;
Best Local Similarity 61.2%  Pred. No. 9,1e-46;
Matches 406; Conservative 0; Mismatches 257; Indels 0; Gaps 0;

```

```

OY 797 CGCTCCATCTCCGGGGCCGCGAGCTGGAGCTGCTTCTGCTGCTGACCGCTCCATGGCG 856
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 186 CGATTGGTCTCGACGACCGCTATGTGAACCATGTGTGGACACCGATCGATGCA 245
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 857 CGGTGTATGGCGGGGCGCTGACGATTACCTGACCCCTGGCCCTCCATGCCAATAG 916
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 246 GAATTCACGCGAGAGTGTCTTAAGCATTTACCTTCACCTTTTTCGGTGGACGCCA 305
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 917 CTGTACAGCCATGTACAGATCGAGAACCATCCGCTGGCGGTGAAGGTGTGTG 976
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 306 TTGTACAAACACCCGACATTCGTATTCAGTTAGCTGTGTGTGTGAAGATCTTGTG 365
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 977 CTAGCGAGACAGAGAGCGCTGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAG 1036
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 366 ATCCAGATGACAGAAAGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 425
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1037 TTTTCAAGTGGCAGACACCAACACAGCTGGGAGATGATGAGAGAGAGAGAGAT 1096
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 426 TTTTCAACTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 485
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1097 GCAGTATCTCTTTACTCGGAGATTTATGTGGCATATTGATGACACCGCTGGGA 1156
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 486 ACAGAAATCTTTTACACGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 545
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1157 AAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1216
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 546 AAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 605
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1217 GGCCTCAGCAGAGCTTCACTGCTGCTCAGAAATCGAATTTACTTGGCTCTCCAT 1276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 606 GGTTTACAAAGCTGCTTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 665
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1277 GACGATTCCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1336
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 666 GATGATGCAAAACAGAGTGCAGAGCTTATGATGAGAACAGAGATTCACATGATGGG 725
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1337 TCCATCTTACAGAGATTTATGATGATGATGATGATGATGATGATGATGATGATG 1396
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 726 TCAATGCTTTTCAACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 785
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1397 ACAGAAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 786 ACATCATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 845
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1457 CTG 1459
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 846 CAG 848
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 6
US-09-369-364A-8
: Sequence 8, Application US/09369364A
: Patent No. 6391610
: GENERAL INFORMATION:
: APPLICANT: Apte, Suneel
: APPLICANT: Hurskainen, Taina L.
: APPLICANT: Hirohata, Satoshi
: TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
: FILE REFERENCE: 26473/4007/10-30-00
: CURRENT APPLICATION NUMBER: US/09/369,364A
: CURRENT FILING DATE: 1999-08-06
: NUMBER OF SEQ ID NOS: 31
: SOFTWARE: PatentIn Ver. 2.1

```

```

: SEQ ID NO 8
: LENGTH: 3638
: TYPE: DNA
: ORGANISM: Mus musculus ADAMTS-8
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (278)..(2992)
: NAME/KEY: misc.feature
: LOCATION: (3636)
: OTHER INFORMATION: n = 7
US-09-369-364A-8

```

```

Query Match          15.6%  Score 237; DB 4; Length 3638;
Best Local Similarity 61.9%  Pred. No. 1,4e-42;
Matches 394; Conservative 0; Mismatches 240; Indels 3; Gaps 1;

```

```

OY 804 TCTCCGGGGCCGCGAGCTGGAGCTGCTTCTGCTGCTGACCGCTCCATGGCGGTGT 863
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 966 TGTCCGAGGCTGCTTCTGCAAAACATTTGTGTGTGTGTGTGTGTGTGTGTGTGT 1025
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 864 ATGGCGGGGCGCTGAGCATTTACCTGACCTGACCTGACCTGACCTGACCTGACCT 923
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1026 ATGGAGCGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1085
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 924 GCGATGTGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 983
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1086 ACCAGCCGAGCATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1145
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 984 ACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1043
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1146 AAGAAAGATGGGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1205
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1044 AAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1103
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1206 GCTGCAAGCGCTTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1265
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1104 TCTCTTTTACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1160
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1266 TCTTGTTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1325
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1161 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1220
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1326 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1385
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1221 TCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1280
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1386 TCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1445
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1281 ATTCCAAATTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1340
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1446 ATTCTAAGCGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1505
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1341 TCCCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1400
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1506 TCTTTCATCCAGCTGAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1565
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1401 AATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1437
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1566 ACCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1602
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 7
US-09-392-184-7/C
: Sequence 7, Application US/09392184
: Patent No. 6395889
: GENERAL INFORMATION:
: APPLICANT: Robison, Keith E.
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN
: FILE REFERENCE: 5800-55
: CURRENT APPLICATION NUMBER: US/09/392,184
: CURRENT FILING DATE: 1999-09-09
: NUMBER OF SEQ ID NOS: 33

```

```

: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 7
: LENGTH: 3126
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(3126)
: OTHER INFORMATION: repolysin (ADAM family of metalloprotease)
US-09-392-184-7

```

```

Query Match          14.0%; Score 211.6; DB 4; Length 3126;
Best Local Similarity 59.7%; Pred. No. 4.3e-37;
Matches 374; Conservative 0; Mismatches 249; Indels 3; Gaps 1;

```

```

QY 797 CGCTCCATCTCCCGGCGCCGAGGTGAGGCTCTGTGTCGTCGACCGCTGCATGCGG 856
DB 2642 CGCTTTGTCTGTGAGCGCGCTTCTGTGAGACGCTGTGTGCGCCGATGCTCATGCT 2583
QY 857 CGTTGTATGGCGCGGCGCTGAGCATTTACTGTACCCCTGCGCTCATGCGCAATAGG 916
DB 2582 GCGTTCTACGGGGCGGCGCGCTTCTGTGAGACGCTGTGTGCGCCGATGCTCATGCT 2523
QY 917 CTGTACAGCCATGCTAGCATTCGAGAACCATCCGCTGCGCTGTGAGGTGCTGTG 976
DB 2522 ATCTACACACCCCGCAGCATCAAGAAATTCATCAACCTGATGTGTAAAGTGTGATC 2463
QY 977 CTAGGGGACAGAGCAAGAGGCTGTGAGTGTGAGAGAGAGTGTGCGACACAGCTCAAGAG 1036
DB 2462 CTAGAGATGAGAAATGGGGCCGAGAGTGTGCGACATGAGGGGCTTACCTGCTGAC 2403
QY 1037 TTTTGCAGTGGCAGCAGCAGACAGACAGCTGGAGATGACCATGAGGAGCATAGCAT 1096
DB 2402 TTCTGCACTGGAGGGGGCTTTCACAGCCGAGGAGGCGGCGCAGACAGCATAGCAG 2343
QY 1097 GCAGCTACTCTGTTTACTGTGGAGAGATTATGTGGCATCA---TTGATGTGACACCTG 1153
DB 2342 ACGGCGATCTCTGTACACAGAGAACTTCTGTGGCAGAGGGGCTGTGTGACACCTG 2283
QY 1154 GGAATGGAGAGCTTGGAGCATATGTTCTCCAGAGCGGAGCTGTGTGATGTAAGAG 1213
DB 2282 GGTGTGGCAGACATCGGACCATTTTGTGACCCACAGAAAGCTGCTGATGATGAGAT 2223
QY 1214 GATGCGCTCCAGCAGCAGCTTCACTGTGCTCAGGAAATGGAGCATTTACTTGGCTTCC 1273
DB 2222 GAGGGGCTCCAGGGCGGCGCAGACCGCTGCGCATGAGGAGCATCTCTCAGCATGCGCC 2163
QY 1274 CATGACGATTCGCAATTTCTGTGAAGAGACCTTTGTTCCAGAGAGATAGAGCTTAATG 1333
DB 2162 CAGCAGACTTCCAAAGCCCTGACACCGCTTTCGGGCGCATGGGCAAGCAGCAGCTGATG 2103
QY 1334 TCTTCATCTCTTACAGCATGATGATGATGATGATGATGATGATGATGATGATGATG 1393
DB 2102 GCACCGCTGTCTGCTCAGCTGAAACGAGCGTGTGCTGCTGCTGCTGCTGCTGCTGAT 2043
QY 1394 ATCAGCAATTCCTGTGATGATGCA 1419
DB 2042 CTCACAGACTTCTTGACGGGGCA 2017

```

```

RESULT 8
US-09-369-364A-3
: Sequence 3, Application US/09369364A
: Patent No. 6391610
: GENERAL INFORMATION:
: APPLICANT: Apte, Suneel
: APPLICANT: Hurskainen, Taina L.
: APPLICANT: Hirahata, Satoshi
: TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
: FILE REFERENCE: 26473/4007/10-30-00
: CURRENT APPLICATION NUMBER: US/09/369,364A
: CURRENT FILING DATE: 1999-08-06
: NUMBER OF SEQ ID NOS: 31

```

```

: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 3
: LENGTH: 1520
: TYPE: DNA
: ORGANISM: homo sapiens ADAMTS-5
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (296)
: OTHER INFORMATION: n-T
: NAME/KEY: misc_feature
: LOCATION: (1443)
: OTHER INFORMATION: n-A
: NAME/KEY: misc_feature
: LOCATION: (1457)
: OTHER INFORMATION: n-G
: NAME/KEY: misc_feature
: LOCATION: (1458)
: OTHER INFORMATION: n-T
US-09-369-364A-3

```

```

Query Match          13.9%; Score 210.2; DB 4; Length 1520;
Best Local Similarity 98.6%; Pred. No. 7.4e-37;
Matches 212; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1253 GCACATTACTTGGCTCTCCCATGACGATTCGCAATTCGTGAGAGACCTTGTGCTCC 1312
DB 1 GCACATTACTTGGCTCTCCCATGACGATTCGCAATTCGTGAGAGACCTTGTGCTCC 60
QY 1313 ACAGAAATGAGCGCTTAAATGCTTCAATCTTACAGATGATGATGATGATGATGATGATG 1372
DB 61 ACAGAAATGAGCGCTTAAATGCTTCAATCTTACAGATGATGATGATGATGATGATGATGATG 120
QY 1373 TCCAAATGACTTTCAGCCACATCAGAGATTCCTGATGATGATGATGATGATGATGATG 1432
DB 121 TCCAAATGACTTTCAGCCACATCAGAGATTCCTGATGATGATGATGATGATGATGATGATG 180
QY 1433 CTGAGCTTACACAGAAAGCAGATCTCTGGCGGGGA 1467
DB 181 CTGAGCTTACACAGAAAGCAGATCTCTGGCGGGGA 215

```

```

RESULT 9
US-09-369-364A-10
: Sequence 10, Application US/09369364A
: Patent No. 6391610
: GENERAL INFORMATION:
: APPLICANT: Apte, Suneel
: APPLICANT: Hurskainen, Taina L.
: APPLICANT: Hirahata, Satoshi
: TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
: FILE REFERENCE: 26473/4007/10-30-00
: CURRENT APPLICATION NUMBER: US/09/369,364A
: CURRENT FILING DATE: 1999-08-06
: NUMBER OF SEQ ID NOS: 31
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 10
: LENGTH: 739
: TYPE: DNA
: ORGANISM: Homo sapiens ADAMTS-8
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (3)...(737)
US-09-369-364A-10

```

```

Query Match          13.5%; Score 205; DB 4; Length 739;
Best Local Similarity 60.4%; Pred. No. 8.5e-36;
Matches 356; Conservative 0; Mismatches 230; Indels 3; Gaps 1;

```

```

QY 797 CGCTCCATCTCCCGGCGCCGAGGTGAGGCTCTGTGTCGTCGACCGCTGCATGCGG 856
DB 60 CGCTTTGTCTGTGAGCGCGCTTCTGTGAGACGCTGTGTGCGCCGATGCTCATGCT 119
QY 857 CGTTGTATGGCGCGGCGCTGAGCATTTACTGTACCCCTGCGCTCATGCGCAATAGG 916

```



Db	4509	CATGAACATGGTGGCTGGCCCTGTTTCATGACCCAGACATTGGAAACCCCATTCACATAC	4450
Qy	958	CGTCGTAAAGSTGCTGTGCTTAGGCACAAGACAGCAAGACCTCGGAAGTGAGCAGAAGACC	1017
Db	4449	CATTGTCCGGCTGGTCTCTGCTGGAMATAGGAGGGAGCCATAAAGATCAGCGACATGC	4390
Qy	1018	TGGCAGCACACTAAAGAACTTTGGCAAGTGGGAGGACCCACACAAACCAAGTGGGAGATGA	1077
Db	4389	AGACAAACACCTGAAGAGCTTTCGCAAGTGGGAGAAAGCATCAACATGAAGGGGAGATGC	4330
Qy	1078	CCATGAGAGACACTAGCATGTACGCTATCCCTGTTTACTCGGGAGATTATGTG-----	1131
Db	4329	CCATCCCTCGACCATGATGACACTGCTATCTCTGCTCAACAGAAAGGACCTGTGTGACGCAT	4270
Qy	1132	GCATCATTCATGTGACACCCCTGGGAATGGCAGACGTTGGGACCAATATGTTCCAGACCG	1191
Db	4289	GAACCGGCCCTGTGAGAACCTCTGGACTGTGCCATGTGGCGGGGCATGTGCAGCGCACCG	4210
Qy	1192	CAGCTGTGCTGTGATTTGAGAGAGATGGCCTCCACGACAGCCTTCACTGTGGCTCACGAAT	1251
Db	4209	CAGCTGCGAGCATACAGGAGACACGGGGCCCTGGCGGTGGCTTCACTGTAGGCCACGAGACT	4150
Qy	1252	CGGACATTTACTGGCCCTCCCATACG	1280
Db	4149	CGGGACACAGTTTGGCACATTCAGCATACG	4121

RESULT 12  
US-09-369-364A-6  
; Sequence 6, Application US/09369364A

```

1  APPLICANT: Apte, Suneel
2  APPLICANT: Hurstainen, Taina L.
3  APPLICANT: Hirohata, Satoshi
4  TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
5  FILE REFERENCE: 26473/4007/10-30-00
6  CURRENT APPLICATION NUMBER: US/09/369,364A
7  CURRENT FILING DATE: 1999-08-06
8  NUMBER OF SEQ ID NOS: 31
9  SOFTWARE: PatentIn Ver. 2.1
10 SEQ ID NO 6
11
12 LENGTH: 3218
13
14 TYPE: DNA
15
16 ORGANISM: Homo sapiens ADAMTS-7
17
18 FEATURE:
19
20 NAME/KEY: CDS
21
22 LOCATION: (13)..(3003)
23
24 US-09-369-364A-6

```

Query Match	8.6%	Score 130.4	DB 4	Length 3218
Best Local Similarity	52.3%	Pred. NO. 1.5e-19		
Matches 368	Conservative 0	Mismatches 321	Indels 15	Gaps 3

[illegible]

QY	1065	AGGGAGGATGACCATGAGACCACTACAGATGACACTTCCCTGTTTCTCGGGAGATT	11244
Db	989	TGAAGGGGATGCCCATCCCTCCACCATGACACTGCCATCTGCTACACAAAGAACCC	10486
QY	1125	TATGTGG-----GCATCATTCATGTGACACCCCTGGGAATGGACAGCTTGGGACATAT	11787
Db	1049	TGTGTGCAGGCATGAAACCGGCCCTGTGAGACCCCTGGGACTGTGCCATGTGGGGGGCATGT	1108
QY	1179	GTTTCCTCAGAGCGCAGCTGTGCTGTGATTTGAAGCAGATGGCTCCACGACGCTTCACTG	12388
Db	1109	GCCACGCCGACCCAGCTGTGCAGCATCAAGAGAGACACGGGCCCTGCGCTGACCTTCACTG	1168
QY	1239	TGGCTACGAAATCGGACATTTACTTGGGCGTCGCCATGACATTTCCAAATTCGTGAAG	12989
Db	1169	TAGGCCACGAGCTTGGGGCAGAGTTTGGCATTCAGATACGGAAAGCGGCAATGACTGTG	12288
QY	1299	AGACCTTGTGGTTCACAGAGATTAAGCGGCTTAATGCTTCCATCCCTTACACGATTTGATG	13588
Db	1229	AGCCCGTTGGGA-----AACGACCTTTTCATCATCTCTCCAAAGCTCCGTATGACAGCGCG	1282
QY	1359	CATCTAAGCCCTGGTCCAAATGCACTTTCAGCCATCAACAGATTCCTGATGATGGCC	14188
Db	1283	CTCCCTCTACACTGGTCCCGCTCGGACGGCGCAGATATATCACAGGTCTTCCTTGACCTGGGT	13422
QY	1419	ATGTGACTGTTTGGCTGAGCATACGACAAAGAGAGATCCCTGGAGC	1462
Db	1343	GGGGCCTGTGCTTGAGACGACCTCTCTGCGCCAAAGACATTTATCGAC	1386

RESULT 13  
US-09-369-364A-14

```

1  APPLICANT:   Apple, Suneeel
2  APPLICANT:   Hurskainen, Taina L.
3  APPLICANT:   Hirohata, Satoshi
4  TITLE OF INVENTION:  Nucleic Acids Encoding Zinc Metalloproteases
5  FILE REFERENCE:  26473/4007/10-30-00
6  CURRENT APPLICATION NUMBER:  US/09/369,364A
7  CURRENT FILING DATE:  1999-08-06
8  NUMBER OF SEQ ID NOS:  31
9  SOFTWARE:  PatentIn Ver. 2.1
10 SEQ ID NO 14
11     LENGTH: 2625
12     TYPE: DNA
13     ORGANISM: Mus musculus ADAMTS-9
14     FEATURE:
15     NAME/KEY: CDS
16     LOCATION: (2)..(2623)
17     US-09-369-364A-14

```

Query Match	8.0%;	Score 121.2;	DB 4;	Length 2625;
Best Local Similarity	-52.68;	Pred. No. 1.4e-17;		
Matches 342;	Conservative 0;	Mismatches 293;	Indels 15;	Gaps 3

[illegible]



QY	1037	TTTTCAAGTGGCAGACACCAACCAACACACAGCTGGGAGATGACCAATGACGAGCAGTACGAT	1096
Db	617	TTTTGCCAGTGGCAGCAGCACTAAAGAACTCTTGGG-----TGGATTTCAGCAGCAGC	667
QY	1097	GCACCTATCCGTGTTTACTCGGGAGATTTATGNGGGCATCATTT---CATGTGACACCCCTG	1153
Db	668	ACAGCCCTGTGGTCAACAAGGGAAGATATCTGGAGAGCTCAGAGACAATGTGACACCTTTA	727
QY	1154	GGAAATGCACAGCCTTGGGACCATATGTTTCTCCAGAGCCGACGCTGTGCTGTGATTGAAGAC	1213
Db	728	GGCTTGCTCAACTGGGAACCATTTGCGACCCCTACCGAAGCTGTTCATTAGTGAAGAC	787
QY	1214	GATGGCCCTCGACGACGCTTCACTGTGGGTCAGCAAGAAATGGACATTTACTGGCCTTCC	1273
Db	788	AGTGGGCTTGAGCACACGCTTTCCAAATAAGCTCACAGCACTGGGCCCATATGTTTAATATGCT	847
QY	1274	CATGACCATTTCCAAATTTCTGTGAAGAGACCTTTGGTTCACAGAAAGATAAAGCGCTTAATG	1333
Db	848	CAGATGTACAGCAATTAATTCACAAGAAAGAAAGAGTTAAGAGTCCCC---AGCATGTCAATG	904
QY	1334	TCCTTCATCCTTACCAAGACATTTGATGCATCTAAGCCCTGCTCCAAATGCACTTACAGCCAC	1393
Db	905	GCACCAACAACCTGAACCTTCTACACCAACCCCTGGATGTGTCTCAAGTCGAGTGGAAATAC	964
QY	1394	ATTCAGCAATTTCTCGATGTATGGCCATGATTACTGTTTSCGTGACCACTTACC	1443
Db	965	ATTCAGTATGTTCTTAGACACTGGGTACGGAGATGCTCTTCTCAATTAAGAAC	1014

```

RESULT 14 1
US-09-392-184-6
Sequence 6, Application US/09392184
Patent No. 6395889
GENERAL INFORMATION:
APPLICANT: Robison, Keith E.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN
TITLE OF INVENTION: PROTEASE HOMOLOGS
FILE REFERENCE: 5800-55
CURRENT APPLICATION NUMBER: US/09/392,184
CURRENT FILING DATE: 1999-09-09
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 6
LENGTH: 703
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(703)
OTHER INFORMATION: repolysyn (ADAM family of metalloprotease)
NAME/KEY: misc_feature
LOCATION: (1)...(703)
OTHER INFORMATION: n = A,T,C or G
US-09-392-184-6

```

Query Match	7.3%;	Score 110.4;	DB 4;	Length 703;
Best Local Similarity	60.8%;	Pred. No. 2.3e-15;		
Matches 180; Conservative	0;	Mismatches 116;	Indels 0;	Gaps 0;

QY 1142 TGTCACACCCCTGGGAATATGCGACAGCGTTGGGACCATTTGTTCTCCAGAGGGCAGCTGTCT 1201

Db 14 TGTGACACCCCTGGGTGTGGGACAGATCGGGACCATTTGTGAGCCCAACAAAGCTGTCTC 73

QY 1202 GTGATTTAAGACGATGGCCCTCCACGCAACCTTCTACTGTGGCTACGAAATGTGGACATTTA 1261

Db 74 GTGATTCAGAGATGAGGGGCTCCAGGGGGCCACACCTCTGGCCCAATGAATAGGCGCAGCTC 133

QY 1262 CTTGAGCCTCTCCCATCGACGATTTCCAAATTTCTGGAAGAGACCTTTGGTTCCACACAGAT 1321

Db 134 CTCAGCATATGGCCCAACGACGACTCCAAAGCCCTGTGACACGGCTCTTTGGGGCCCATATGGGCAAG 193

QY 1322 AAGGCGTTAATGTCTTCATCTTTACCAAGCATTTGATGCATCTAAGCCCTGGTCCAAATGC 1381

Db 194 CACACAGTGGACCGCGTTCGTCACCTGTAACCAACGCGTGGCGTCCCGTC 253

Qy 1382 ACTTCAGGCACACATCACAAGATTCCTGGATGATGGCCATGTTACTGTTGCTGGA 1437

Db 254 AGCGCCATGTATCTCACAGAGCTTCGGACGGGGGCGACAGAGACTGTCCTCGGA 309

```

RESULT 15
US-09-369-364A-12
; Sequence 12, Application US/09369364A
; Patent No. 6591610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12

```

```

? ORGANISM: Homo sapiens ADAMTS-9
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (3)..(5648)
? NAME/KEY: misc_feature
? LOCATION: (1405)
? OTHER INFORMATION: n-T
? NAME/KEY: misc_feature
? LOCATION: (1563)
? OTHER INFORMATION: n-T
? US-09-369-364A-12

```

Query Match	6.38;	Score 95.8;	DB 4;	Length 5804;
Best Local Similarity	49.18;	Pred. No. 5.2e-12;		
Matches 318;	Conservative	0;	Mismatches 317;	Indels 12; Gaps 2;

Oy	797	CGCTCAATCTCCCGGGCCCGCAGAGTGAGCTCTCTTGGTGTACAGCGTCCATGGG	856
Dd	705	CGTTTTTATTCCTATCCACGGTTTGTGAAGTCTTGTGTGTGCAGACAACGAATGTT	764
Oy	857	CGGTTTGAATGCGCGGGCCCTGACGACATTTACCTGCTGACCCCTGGCTCCATGCAATAG	916
Dd	765	TCATATCCATGAGAAACCTTCAACATAATTTTAACTTAATGTCAATGTAGCTCT	824
Oy	917	CTGTACAGCCATGCTAGCATTCGAGAACCAATCCGCTTGCCCTGGTGAAGTGGTGTG	976
Dd	825	ATCTATTAAGCCCAACAGCTATTGGAAATTAATTAATTTGTATTTGTGAACCTTAATTTGG	884
Oy	977	CTAGCGCAGAGCAGAGAGCCTGGAAGTAGAGAACAGACCTGCCACCACTCCAAAGAC	1036
Dd	885	ATTCATATATGAAAGAGATGAGGCCCTTCATATCTTTTAATGTCATGACACACATTTAAAAAC	944
Oy	1037	TTTTTGCAAGTGTGGCAGCACAACACACCACTGGGAGATGACCATAGAGACCACTACAT	1099
Dd	945	TTTTTGCAGTGGCAGCACTTGCAGAAAGTCACGGGTATCCATCATATATAC -----T	995
Oy	1097	GCAGCTATCCGTTTACTCGGAGAGATTTATGTGGCAGCATTCATCATAGTACACCCGGGA	1156
Dd	996	GCTGTTCTCTTAACACAGACAGGATATCTGGAGAGCTACGACCAAAATGTATACCTTAGCC	1055
Oy	1157	ATGGCAGACGTTGGGACCATATATGTTCTCCAGAGCCGACGCTGTCTGTATGAAGCAT	1216
Dd	1056	CTGGCTGAACGGGGAACCACTTTTGATATCCCTAATAGAAGCGTCTATTGTGTAAGATATGT	1115
Oy	1217	GGCCCTCCAGCAGGCTTCACATGTTGGCTCAGACGAATGGAGATTTACTTGGCCCTTCACAT	1276
Dd	1116	GGATTGAGTACAGCTTTTACGATGCCCAATGAGACTGGCGCATGTGTTAACTAGCTCAT	1175
Oy	1277	GACGATTTCCAAATTTCTGTGAAGACCTTTTGTTCCACAGAGATTAAGCGCTTAATGTCT	1336

```
Db 1176 GATGACAAACAAATGCTAAAGA---AGAAGAGTTAAGAGTCCCAGCATGTGATGGCT 1232
Qy 1337 TCCATCCTTACCGACATGTGATGCTAAGCCCTGGTCCAAATGCACCTTCAGCCACCATC 1396
Db 1233 CCACACCTGAACCTTCTACACCAACCCCTGGATGTGTCAAAGTGTAGTCGAAAAATATATC 1292
Qy 1397 ACAGATTCTCGATGATGCGCCATGCTGTTGCTGACCTAAC 1443
Db 1293 ACTGAGTTTCTAGACACACTGGTTATGGCGAGTGTTCCTTAACGAACC 1339
```

Search completed: April 26, 2003, 05:36:28  
Job time : 128.77 secs

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 25, 2003, 07:28:33 ; Search time 664.931 Seconds  
(without alignments)  
5134.413 Million cell updates/sec

Title: US-10-050-200-2

Perfect score: 1516  
Sequence: 1 gtcagcgcagcgcactatgc.....agcttgcgagaagtag 1516

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_101002.\*

```
1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	1453	95.8	3250	20	AAx00726 Human aggrecan deg
2	1427	94.1	5530	21	AAc75435 Human ORF990
3	1010.8	66.7	2751	21	AAa95822 Rat metalloprotein
4	1010.6	66.7	3002	22	AAa63437 Murine ADAMTS-5 CD
5	569.8	37.6	961	24	ABO48540 Oligonucleotide fo
6	569.8	37.6	961	24	ABO48541 Oligonucleotide fo
7	561.2	37.0	961	24	ABO48538 Oligonucleotide fo
8	561.2	37.0	961	24	ABO48539 Oligonucleotide fo
9	517.2	34.1	957	24	ABO48542 Oligonucleotide fo

10	517.2	34.1	957	24	ABO48543	Oligonucleotide fo
11	507.6	33.5	957	24	ABO48544	Oligonucleotide fo
12	507.6	33.5	957	24	ABO48545	Oligonucleotide fo
C	425	28.0	528	22	AAc90360	AA286689 CDNA clon
13	267	17.6	2853	22	AAH20224	Human ADAM-lype me
14	267	17.6	3261	20	AA232000	Human METH1 encodi
15	267	17.6	3261	22	AAc90057	Human integrin lig
16	267	17.6	4014	19	AAV65008	Human Tingo-71 enc
17	267	17.6	4676	20	AAx19955	Human secreted pro
18	267	17.5	4712	20	AAx04374	Human METH1 relate
19	265.8	17.5	4112	20	AAc34374	Human metalloprote
20	258.6	17.1	4180	20	AA232010	D67076 CDNA clone.
21	258.6	17.1	4180	22	AAc90067	Human metalloprote
22	257.8	17.0	2853	22	AAH41003	Human protease CDN
23	257.8	17.0	2853	24	AAc35569	Human protease PR
24	257.8	17.0	2930	24	ABx12894	Human metalloprote
25	257.8	17.0	2937	24	AAc97182	Human metalloprote
26	257.8	17.0	3446	24	AAc35571	Human protease CDN
27	256	16.9	4858	24	ABN86294	Human ADAMTS-1 cod
28	253	16.7	2184	20	AAx17990	Rat metalloprotein
29	246.6	16.3	1518	21	AAa55827	Murine ADAMTS-8 CD
30	237	15.6	3638	22	AAc3441	Human METH2 encodi
31	223.2	14.7	3008	20	AA232001	Human METH2 coding
32	223.2	14.7	3008	22	AAc90056	Human ADAM-lype me
33	221.6	14.6	2346	21	AAa95821	Human metalloprote
34	221.6	14.6	2670	22	AAH20226	Human ADAM-lype me
35	221.6	14.6	2670	22	AAc92166	Human ADAM-lype me
36	211.6	14.0	3126	24	ABN86300	Human protease CDN
37	210.2	13.9	1520	22	AAc3438	Human ADAMTS-5 CDN
38	206.6	13.6	1143	21	AAa55831	Human metalloprote
39	205	13.5	739	22	AAc3442	Human ADAMTS-8 CDN
40	180.6	11.9	1104	24	AAc35570	Human protease CDN
41	175.8	11.6	4192	20	AAx00722	Human aggrecan deg
42	175.8	11.6	4301	22	AAa58146	Human aggrecanase-
43	175.8	11.6	4303	21	AAa5826	Human metalloprote
44	174.4	11.5	652	22	AAc90338	AI400905 CDNA clon
45	174.2	11.5	4406	22	AAa46100	Human DNA encoding

#### ALIGNMENTS

```
RESULT 1
AAx00726 standard; CDNA to mRNA; 3250 BP.
ID      AAX00726;
AC      AAX00726;
DT      30-MAR-1999 (first entry)
XX      Human aggrecan degrading metalloprotease 2 gene.
DE      Human; aggrecan degrading metalloprotease; cartilage; proteoglycan;
XX      interglobular domain; matrix metalloprotease; bovine; interleukin-1beta;
KW      primer; PCR; amplification; inhibitor; cleavage; inhibitor; ADMP; ss;
KW      arthritis; joint injury; pseudogout.
OS      Homo sapiens.
XX      Homo sapiens.
FH      key
FT      CDS
FT      121..2913
FT      Location/Qualifiers
FT      /*tag=a
FT      /product= ADMP-2
FT      /note="aggrecan degrading metalloprotease 2"
W09905291-A2.
04-FEB-1999.
24-JUL-1998; 98MO-US15438.
16-OCT-1997; 97US-0062169.
25-JUL-1997; 97US-0053850.
15-AUG-1997; 97US-0055836.
```

XX	PA	(DUPPO ) DU PONT PHARM CO.
XX	P1	Arner EC, Burn TC, Copeland RA, Decicco CP, Liu R, Magolda R;
XX	P1	Pirata M, Solomon KA, Tortorella MD, Tizaskos JM, Yang F;
XX	DR	WPI; 1999-142943/12.
XX	DR	P-PSDB; AAM75426.
PT	PT	New Isolated aggrecan degrading metallo proteases - used to develop
PT	PT	products for treating, e.g. osteoarthritis, joint injury, resective
xx	xx	arthritis, psoriatic arthritis or juvenile rheumatoid arthritis
PS	PS	Claim 17; Page 65-68; 73pp; English.
XX	CC	This sequence represents the coding region for the human aggrecan
CC	CC	degrading metalloprotease 2 (ADMP-2). ADMP-1 (AAK00725) and ADMP-2 are
CC	CC	novel proteases that cleave the aggrecan (a major cartilage
CC	CC	proteoglycan) between residues glu373-Ala374 of the interglobular domain
CC	CC	(compared with cleavage between Asn341-Phe342 by the matrix
CC	CC	metalloproteinases MMP-1, -2, -3, -7, -8, -9 and -13). ADMP-1 and ADMP-2
CC	CC	were isolated and purified from the conditioned media of bovine nasal
CC	CC	cartilage stimulated by Interleukin-1beta. The purified proteins were
CC	CC	partially sequenced and primers were synthesised based on the resultant
CC	CC	amino acid sequences (AAX00727-X00732 for ADMP-1 and AAX00733-X00736 for
CC	CC	ADMP-2). These were used to isolate the corresponding genes from cDNA.
CC	CC	The ADMP polypeptides can be used for identifying inhibitors of ADMP
CC	CC	activity which would prevent cleavage of the aggrecan core protein,
CC	CC	thereby decreasing the loss of aggrecan from cartilage. Such inhibitors
CC	CC	can be used for treating diseases such as osteoarthritis, joint injury
CC	CC	reactive arthritis, acute pyrophosphate arthritis (pseudogout), psoriatic
CC	CC	arthritis and juvenile rheumatoid arthritis.
XX	SQ	Sequence 3250 BP; 761 A; 900 C; 919 G; 670 T; 0 other;
		Query Match 95.8%; Score 1453; DB 20; Length 3250;
		Best local similarity 99.7%; Pred. No. 7.1e-307;
		Matches 1456; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY	7	GCAGCGCACTATGCTGCTCGGGTGGGCGTCCTGCTGTGTCGGGGTTCCGGCTGCCCT 66
Db	111	GCAGCGCACTATGCTGCTCGGGTGGGCGTCCTGCTGTGTCGGGGTTCCGGCTGCCCT 170
QY	67	GCCCCCGGTGGCCCCCGCCGACACACTTGCCAGATTAAGCCGGGGAACCTTCCGACTGC 126
Db	171	GCCCCGGTGGCCCCCGCCGACACACTTGCCAGATTAAGCCGGGGAACCTTCCGACTGC 230
QY	127	TGCAGCACC GCCCGCACGCCCGCGCGCGAGGGGAGAAGTAGTGCAGAGAGACCGAGCC 186
Db	231	TGCAGCACC GCCCGCACGCCCGCGCGCGAGGGGAGAAGTAGTGCAGAGACCGAGCC 290
QY	187	TCCCGGCCACCCGCAACCCCCTTGCGCAGCGGCGCAGAGAACAAAGGGCTGTGCAGAACT 246
Db	291	TCCCGGCCACCCGCAACCCCCTTGCGCAGCGGCGCAGAGAACAAAGGGCTGTGCAGAACT 350
QY	247	GCACAACACTTACTCTCGGCGGGGCGGAAGGTGGGTACTGCTGATAGCGGGGGCGCGAG 306
Db	351	GCACAACACTTACTCTCGGCGGGGCGGAAGGTGGGTACTGCTGATAGCGGGGGCGCGAG 410
QY	307	GTTCTCTTGGACCTTGAGCGAGATGGTGGTGGGCAATTGCTGGCTTGTGCCCGCAGG 366
Db	411	GTTCTCTTGGACCTTGAGCGAGATGGTGGTGGGCAATTGCTGGCTTGTGCCCGCAGG 470
QY	367	AGCGGGAGAGATGGGCCCTTGCGCGCCACCGAGCCACACTGCTTATCGGGGACAGTGA 426
Db	471	AGCGGGAGAGATGGGCCCTTGCGCGCCACCGAGCCACACTGCTTATCGGGGACAGTGA 530
QY	427	CGGTAGTATCCCGCTTGTGGCTGTGTTGACTCTGTGGGGGTCTGCAGAGGCTTCTTCCG 486
Db	531	CGGTAGTATCCCGCTTGTGGCTGTGTTGACTCTGTGGGGGTCTGCAGAGGCTTCTTCCG 590
QY	487	GGTCAAGCAGCGGCGCTACACCTTAAGCCACTGCTGCGCGGAGACCTTGGGCGAGAGAGA 546

Db	591	GGCAAGCAGCGCGGCTACACACCTAAAGGACACTGCTGGCGGAGACCTTGGGCGGAGGAAGA	650
QY	547	AAAGGGCGCGCTGTACGGGGATGGGTCCGACGGATCTCTGCACGCTCTACACCCGCGAGGG	606
Db	651	AAAGGGGGCGCTGTACGGGGATGGGTCCGACGGATCTCTGCACGCTCTACACCCGCGAGGG	710
QY	607	CTTCAGCTTCAGGCGCCTTGCGCGCGCGCGCGCAGCTCTCGAATAACCCCGGGTCCACACCGGA	666
Db	711	CTTCAGCTTCAGGCGCCTTGCGCGCGCGCGCGCAGCTCTCGAATAACCCCGGGTCCACACCGGA	770
QY	667	GGCCCGACGAGCATGCTCCGGCGGACAGCAACCCGAGCGGAGCGCGAGCACTGGCCTCGCA	726
Db	771	GGCCCGACGAGCATGCTCCGGCGGACAGCAACCCGAGCGGAGCGCGAGCACTGGCCTCGCA	830
QY	727	GCTCTTGAGACAGTCCGCTCTCTCGCCCGCTGGGGGCTAGAACCGGAGAGTGGTGGCG	786
Db	831	GCTCTTGAGACAGTCCGCTCTCTCGCCCGCTGGGGGCTAGGAACCGGAGAGTGGTGGCG	890
QY	787	GGCGGGCGCGGCTGCATCTCCCGGGCGCGCAGGTGAGAGTCTCTTGTGTGGCTGACGC	846
Db	891	GGCGGGCGCGGCTGCATCTCCCGGGCGCGCAGGTGAGAGTCTCTTGTGTGGCTGACGC	950
QY	847	GTCCATGGCGCGGTTGTATATGGCGGGGCGCTGCAGCATTTACCTGTGACCCTGGCTCCAT	906
Db	951	GTCCATGGCGCGGTTGTATATGGCGGGGCGCTGCAGCATTTACCTGTGACCCTGGCTCCAT	1010
QY	907	CGCCATATAGGCTGTACAGGACATGCTATAGCATCGAGAACACATCCGCTGGCGTGGTAA	966
Db	1011	CGCCATATAGGCTGTACAGGACATGCTATAGCATCGAGAACACATCCGCTGGCGTGGTAA	1070
QY	967	GGTGTGTGCTGTAGCGCAAGAGCAAGAGCCTTGGAGTGTAGACAAAGACCTGCGCACAC	1026
Db	1071	GGTGTGTGCTGTAGCGCAAGAGCAAGAGCCTTGGAGTGTAGACAAAGACCTGCGCACAC	1130
QY	1027	ACTCAAGAACTTTGGCAAGTGGCAGACCAACCAACAGCTGGGAGATGACATGAGGA	1086
Db	1131	ACTCAAGAACTTTGGCAAGTGGCAGACCAACCAACAGCTGGGAGATGACATGAGGA	1190
QY	1087	GCACCTAGCATGACGCTATCTGTTTACTCGGGGAGATTTATGTGGGCATCATTCATGTGA	1146
Db	1191	GCACCTAGCATGACGCTATCTGTTTACTCGGGGAGATTTATGTGGGCATCATTCATGTGA	1250
QY	1147	CACCTGGGAATGGCAGACGTTGGAGACATATCTTCCAGAGGCGACCTGTGCTGAT	1206
Db	1251	CACCTGGGAATGGCAGACGTTGGAGACATATCTTCCAGAGGCGACCTGTGCTGAT	1310
QY	1207	TGAAGACGATGGCCTCCAGCGCAGCTCTACTGTGGCTCAGCAATTCGAGCAATTTACTTGG	1266
Db	1311	TGAAGACGATGGCCTCCAGCGCAGCTCTACTGTGGCTCAGCAATTCGAGCAATTTACTTGG	1370
QY	1267	CCTCTCCATGACGATTCGAAATTCGTGTGAAGAGACCTTTGGTCCACAGAAAGTAAAGG	1326
Db	1371	CCTCTCCATGACGATTCGAAATTCGTGTGAAGAGACCTTTGGTCCACAGAAAGTAAAGG	1430
QY	1327	CTTAATGCTTCCATCTCTTACCAAGCAATTTGATGATCTAAGACCTGTGCAAAATGACATTC	1386
Db	1431	CTTAATGCTTCCATCTCTTACCAAGCAATTTGATGATCTAAGACCTGTGCAAAATGACATTC	1490
QY	1387	AGCCACATCACAGATTTCTGTGATGATGGCCATGTAATCTTTGTCTGACCTACACAG	1446
Db	1491	AGCCACATCACAGATTTCTGTGATGATGGCCATGTAATCTTTGTCTGACCTACACAG	1550
QY	1447	AAAGCAGATCTCTGGGCGGGA	1467
Db	1551	AAAGCAGATCTCTGGGCGGGA	1571
RESULT 2			
AAC75435			
AAC75435 standard: cDNA; 5530 BP.			
AC AAC75435;			





```

OY 772 GCAGAGCTGGTGGGCGGCGCGCGCTCCATCTCCGCGCGCGCGCGAGGTGAGTGTCT 831
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 892 TCAGACTGTGTGAGGAGGCGCGCTTCATCTCCAGCGCGCGCGAGGTGAGTGTCT 951
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 832 TCTGTGGCTGACGGCTCCATGGCGGTTGTATGGCCCGCGCGCTGACACATTACTGCT 891
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 952 CTGTGGGCTGACTCTTCCTGCGCAAGATGTATGGCGCGCGCTGACACATTACTGCT 1011
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 892 GACCCGTGGCTTCATGCGCAATAGGCTGTACAGCCATGTACATGAGAACACATCCG 951
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1012 GACCTCGGCTCTATTGTCCACACCGGCTGTACATGTACATGAGAACACATCCG 1071
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 952 CCTGGCGGCTGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1011
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1072 CTTGGCGGCTGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1125
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1012 GACCGCTGGCACACACTTAAGACTTTTGCAGAGTGGCAGCAGCAACACACACAGTGG 1071
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1126 GACCGCGGCGCAGCGCTTCAGAACTTTTGCAGAGTGGCAGCAGCAACACACAGTGG 1185
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1072 ACATGACATGAGAGACATGACATGACATGACATGACATGACATGACATGACATGAC 1131
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1186 TCATGACATGAGAGACATGACATGACATGACATGACATGACATGACATGACATGAC 1245
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1132 GCATCATTCATGTGACACCGCTGGGAATGGCAGACGTTGGGACCATATGTTCTCCAGACG 1191
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1246 GCATCATTCATGTGACACCGCTGGGAATGGCAGACGTTGGGACCATATGTTCTCCAGACG 1305
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1192 CAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1251
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1306 CAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1365
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1252 CGACATTTACTTGGGCTCTCCATGACATGACATGACATGACATGACATGACATGAC 1311
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1366 TGCACATCTACTTGGGCTCTCCAGACATGACATGACATGACATGACATGACATGAC 1425
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1312 CACAGAGATTAAGCGCTTAATGTCTTCATGCTTACACAGATTAATGATGATGATGATG 1371
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1426 TACAGAGATTAAGCGCTTAATGTCTTCATGCTTACACAGATTAATGATGATGATGATG 1485
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1372 GTCCAAATGCTTACGACCATGACCATGACCATGACCATGACCATGACCATGACCATG 1431
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1486 GTCCAAATGCTTACGACCATGACCATGACCATGACCATGACCATGACCATGACCATG 1545
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1432 GCTGACCTACCAAGAAAGATGCTGGGCGGGA 1467
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1546 ACTAGATGTACCAAGAAAGATGCTGGGCGGGA 1581
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 4  
AA63437  
ID AAF63437 standard; cDNA; 3002 BP.

```

XX AC AAF63437;
XX DT 14-MAY-2001 (first entry)
XX DE Murine ADAMTS-5 cDNA sequence.
XX ADAMTS-N; disintegrin; metalloprotease; thrombospondin type I motif;
XX tumour cachexia; inflammation; dermatosparaxis; EDS-VIIC; angiogenesis;
XX Ehlers-Danlos syndrome type VIIC; articular cartilage erosion; mouse;
XX metastasis; embryogenesis; egg implantation; chromosome 16; ADAMTS-5; ss.
OS Mus musculus.
XX WO20011074-A2.
XX 15-FEB-2001.
XX PD 03-AUG-2000; 2000MO-US21223.
XX PF 06-AUG-1999; 99US-0369364.
XX PR

```

```

XX (CLEV-) CLEVELAND CLINIC FOUND.
PA (APTE/) APTE S. S.
PA (HURS/) HURSKAINEN T. L.
PA (HIRO/) HIROHATA S.
XX Apte SS, Hurskainen TL, Hirohata S;
XX WPI: 2001-159978/16.
XX P-PSDB: AAB72280.
XX
XX Murine and human 'A Disintegrin-like And Metalloprotease domain with
XX Thrombospondin type I motifs' proteins and the nucleic acids encoding
XX them, useful for treating e.g. tumours, inflammation and arthritis -
XX
XX Claim 15; Fig 1; 181pp; English.
XX
XX This invention relates to murine and human ADAMTS-N (A disintegrin-like
XX and metalloprotease domain with thrombospondin type I motifs) proteins,
XX designated ADAMTS-5, 6, 7, 8, 9, 10 and R1. Also included in the
XX invention are cDNA sequences encoding the proteins, and antibodies
XX specific for the proteins. The nucleic acid sequences and proteins may be
XX used in the prevention, diagnosis and treatment of diseases associated
XX with inappropriate ADAMTS-N expression. Disorders that may be treated
XX using the nucleic acids, proteins and antibodies include, for example
XX tumour cachexia, inflammation, dermatosparaxis in cattle or Ehlers-Danlos
XX syndrome type VIIC (EDS-VIIC) in humans, erosion of articular cartilage
XX in arthritic (both inflammatory and non-inflammatory) disease,
XX angiogenesis, tumour growth and metastases, and they may also be used for
XX controlling embryogenesis and implantation of fertilised eggs. The
XX present sequence represents cDNA encoding murine ADAMTS-5. The murine
XX ADAMTS-5 gene is located on chromosome 16.
XX
XX Sequence 3002 BP: 727 A; 787 C; 845 G; 643 T; 0 other:
XX
XX Query Match 66.7%; Score 1010.6; DB 22; Length 3002;
XX Best Local Similarity 82.7%; Pred. No. 1.5e-210;
XX Matches 1218; Conservative 0; Mismatches 219; Indels 36; Gaps 4;
XX
XX 7 GCAGCGCACTATGCTGCTGGGCGGCGCTC-----CTGCTGCTGGCGC 51
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 8 GCAGCGCACTATGCTGCTGGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 67
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 52 GTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 111
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 68 GTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 127
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 112 GCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 171
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 128 GCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 187
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 172 GAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 231
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 188 GAGCGCA-----GCGCATTTACAACTCTTGGCGGCGGCGGCGGCGGCGGCGGCGG 235
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 232 GCTGTGTGAGAACATGAGCACTACTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 291
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 236 GCTGTGTGAGAACATGAGCACTACTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 295
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 292 CCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 350
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 296 CCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 355
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 351 -GCTTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 408
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 356 TACATGCTTACTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 415
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 409 CTATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 468
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 416 CTACAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 475
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 469 TCTGAGCGGCTTCTTGGCGGTCAGGCGGCGGCTACACCTTAAGCCACTGCTGCGCGG 528
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```





```

Db 761 CCCGACACCCCGCCCTTAACGACAGACGCAAAAACAAAAAATAATACAAACATC 702
Oy 248 GACCAACTCTACTCCGCGCGCGGCAAGTGGCTACTCTGTCTAGCGCGCGCGCAG 307
Db 701 GACCAACTCTACTCCGCGCGCGGCAAAATAAATACTACTCTGTCTAGCGCGCGGAAA 642
Oy 308 TTCCTCTTGACCTTGAGCGAGATGTTGGTGGGCTTGTGCTTCTGCTCCGAGGA 367
Db 641 TTCCTCTTAACCTTAACGAAATATTCGATTAACATTACTACTCTGTACCCGAAAA 582
Oy 368 GCGGAGAGAGTGGCCCTGGCGCCACCGGAGCCAGTCTTATCTAGCGCGCGCAGGAG 427
Db 581 AACGAAAGCAATATGCGCTTACCGCACGAAACCTACTCTTATGGAACCAATAAAC 522
Oy 428 GGTAGTCCCGCTCTCTGTGCTCTTTGACCTCTGTGGGGTCTGACGCGCTTCTCGCG 487
Db 521 GCTAATCCCGCTCTCTACTACTTTTAACTCTTAAATAATCTGAGACACTTCTCGCG 462
Oy 488 GTCAAGCAGCGCGCTTACACCTTAAGCAGCTGTGCGCGGACCTGGCGGAGAGAA 547
Db 461 ATCAAAACAGCGCGCTTACACCTTAAGCAGCTGTGCGGAGCCTTAAGCAAAAAAAA 402
Oy 548 AAGGGCGCGGTACGCGGATGGTCCGACAGATCTGACGTCTACACCGCGGAGGCG 607
Db 401 AAAAAACGCTTACGAAATAATCCGACGAAATCTTACAGCTTACACCGCGGAAAC 342
Oy 608 TTACGCTTTCAGGCGCGCTGCGCGCGCGGCGGAGTGGGAAACCGCGCTCAACCGGAG 667
Db 341 TTCACTTTAAACCTTACCGCGCGCGCGGCGGCACTACGAAACCGCGGCTTCAACGGA 282
Oy 668 GCGCAGAGATCTCTCGCGCGGCGGAGCAACCGGAGCGCGGAGCACTGGCTTGGAG 727
Db 281 ACCCAGCAATATTCGCGGAGCAACCAACCGGAGCGGCGGCACTTAACTGCGAA 222
Oy 728 CTCTTGAGCACTGCGCTCTGCGCGCGCGGCTGAGAGCGGAGCGGTGGGCGG 787
Db 221 CTTTAAACCTTACCGCTCTTTCGCGCGCTTAAACCTCAAAACCGGAAACGTAATACGA 162
Oy 788 GCGGCGCGCGCTCCATCTCCGCGCGCGGCGGAGGTGAGCTCTTCTGTGCTGAGCGG 847
Db 161 CCACGACGCGCTCCATCTCCGCGCGGAAACCGGCAAACTTCTTAACTAACTAACGCG 102
Oy 848 TTCATGCGCGGTGTGATGCGCGGCGGCTGAGCACTTACTCTGTGACCTTGGCTTCATC 907
Db 101 TCCATTAAGCGGATTAATTAACGAAACCTTAACCACTTACTTAACCTTACCTTCATC 42
Oy 908 GCCAATAGGCTGACAGCCATGCTAGCATGAGAACCAAT 948
Db 41 GCCAATTAACCTTACACCATCTACTAATCATGAAACCAAT 1

```

```

XX 01-SEP-2000; 2000DE-1043826.
PR 05-SEP-2000; 2000DE-1044543.
XX
PA (EPIC-) EPIGENOMICS AG.
XX Olek A, Piepenbrock C, Berlin K, Guetig D;
DR WPI; 2002-371829/40.
XX
PT Determining the degree of cytosine methylation in genomic DNA, useful
PT for diagnosis and prognosis, comprises selective hybridization of
PT amplicons from chemically treated DNA
XX
PS Claim 12; 56bp + Sequence Listing; 56bp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one
CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
CC and the degree of hybridisation to both classes is determined from the
CC label on the amplicon. From the ratio of labels hybridised to the two
CC classes of oligomers, the degree of methylation is calculated. The method
CC is used: (i) for diagnosis and/or prognosis of side effects of
CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
CC of the central nervous, cardiovascular, gastrointestinal and respiratory
CC systems etc., particularly by detecting mutations or single nucleotide
CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
CC types and for investigating cell differentiation. The method allows the
CC methylation status of many C residues to be determined simultaneously.
CC AB013410-AB054121 represent genomic DNA sequences used to illustrate the
CC method for determining the degree of cytosine methylation described in
CC the disclosure of the invention.
XX
SQ Sequence 961 BP; 367 A; 341 C; 102 G; 151 T; 0 other;
XX
Query Match 37.6%; Score 569.8; DB 24; Length 961;
Best local similarity 75.3%; Pred. No. 1,1e-114;
Matches 709; Conservative 0; Mismatches 232; Indels 0; Gaps 0;

```

Db	441	GCTATATCCCGCTCTCTAATATCTTTAACTTAAATATCGACGACTTCTTCGC	500
Qy	488	GTCGAAGCGGGGCGCTTACACCTTAAACCACTGGTGGGCGGAGTGGGGAGAGA	547
Db	501	ATTCAAACACGGGGCGCTACACCTTAAACCACTACTACGCCACCTTAAAGAAAA	560
Qy	548	AAGGGCGCGCTGTACGGGGATGGGTCCGACGATCTCTGCAGCTTACACCGCGAGGC	607
Db	561	AAAAAAGCGGTATACGAAATAATTAATTCGACAGAAATCCTACACGCTACACCGGAAAA	620
Qy	608	TTTCAGCTTCGAGGCGCTTCCGCGCGCGCCAGCTGCGCAAAACCCCGCTCACACGGAG	667
Db	621	TTTCAACTTTAAACACCTTACCGCGCGCGCGCAACTACAAAAACCCCGCTCACACGAAA	680
Qy	668	GCCCAAGCAGCATGCTCCGGGCGCAGACAAACCCGAGCGGAGCAGTGGCTCGCAG	727
Db	661	ACCCACGACACATATTCGAGCGACAAACCCGACGACGCGGACATTAACCTCGCA	740
Qy	728	CTCTTGACACAGTCCGCTCTCTCCGCGCTGGGGGCTCAGAGACCGACGAGTGGTGGCG	787
Db	741	CTTTTAAACCAATCCGCTCTTTCGCGCGCTTAAACCAAAACCGCAACGTAATTAACGA	800
Qy	788	CGCGCGCGCGCGCTCCATCTCTCCGCGCGCGCGAGTGSAGCTGCTTCGGTGGCTGACGC	847
Db	801	CGACGACCGCGCTCATCTCTCCGAACCGCGCAATTAACCTATTAACTTAACCTACGC	860
Qy	848	TCCATGCGCGGTTGTATGGCGGGGCGCTGCGAGCATTAACCTGGTGAACCTCGCTCCATC	907
Db	861	TCCATTAACGGCATATATATTAACCGAAACCTAACACATTACTACTTAACCTTACCTGATC	920
Qy	908	GCCCAATAGCTGTACAGCCATGCTAGCATGAGAACCAACAT	948
Db	921	GCCCAATAACTATACCAACCATATACATCGAAAAACCAT	961
RESULT 7			
ABQ48538			
ID	ABQ48538 standard; DNA: 961 BP.		
XX	ABQ48538:		
AC	12-JUL-2002 (first entry)		
DT	Oligonucleotide for detecting cytosine methylation SEQ ID NO 35129.		
DE			
XX	Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;		
KW	drug; side effect; cancer; central nervous system; cardiovascular;		
KW	gastrointestinal; respiratory system; single nucleotide polymorphism;		
KW	SNP; cell differentiation; ds.		
XX	Homo sapiens.		
OS	MO2020218632-A2.		
XX	07-MAR-2002.		
PN	01-SEP-2001; 2001MO-EPI0074.		
XX	01-SEP-2000; 2000DE-1043826.		
PR	05-SEP-2000; 2000DE-1044543.		
XX	(EPIG-) EPIGENOMICS AC.		
XX	Olek A, Piepenbrock C, Berlin K, Guelzig D;		
PI	WPI; 2002-371829/40.		
DR	Determining the degree of cytosine methylation in genomic DNA, useful		
PT	for diagnosis and prognosis, comprises selective hybridization of		
PT	amplicons from chemically treated DNA		
XX	Claim 12; 56bp + Sequence Listing; 56bp; German.		

20 This invention describes a novel method for determining the degree of  
21 methylation of a particular cytosine in a motif 5'-CGC-3', present in a  
22 genomic sample of DNA. The sample is treated chemically to convert  
23 cytosine (C) but not methylated C, to uracil, then part of the genomic  
24 DNA that contains the target C is amplified to form a labeled amplicon.  
25 The amplicon is hybridised to two classes, each with at least one  
26 member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers  
27 and the degree of hybridisation to both classes is determined from the  
28 label on the amplicon. From the ratio of labels hybridised to the two  
29 classes of oligomers, the degree of methylation is calculated. The method  
30 is used: (1) for diagnosis and/or prognosis of side effects of  
31 therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders  
32 of the central nervous, cardiovascular, gastrointestinal and respiratory  
33 systems etc., particularly by detecting mutations or single nucleotide  
34 polymorphisms (SNP's); and (11) for differentiation of cell or tissue  
35 types and for investigating cell differentiation. The method allows the  
36 methylation status of many C residues to be determined simultaneously.  
37 ABO13110-AB054121 represent genomic DNA sequences used to illustrate the  
38 method for determining the degree of cytosine methylation described in  
39 the disclosure of the invention.

Sequence 961 BP; 138 A; 102 C; 331 G; 390 T; 0 other;

Query Match	37.08;	Score 561.2;	DB 24;	Length 961;
-------------	--------	--------------	--------	-------------

Matches 704; Conservative 0; Mismatches 238; Indels 0; Gaps 0.

QY	7	GCACGACACATTCGTGCTCCGGGTGGGGGTCCCTCTCTCTGTGGCCGTTCCGCTGCCCT	66
Dp	20	GTAGCGTATATATGTTGTTTCGGGTGGGGGTTTGTGTGTGGCCGCTTTCGTTTGT	79
QY	67	GGCCGCGGTGGCCCCCGCGCGACACTCTCCAGATTAAGCCGGGAGCGCTCCGACTGC	126
Dp	80	GGTCGCGGTCCGTTTCGTCGGATATTTGTTTACGATAAAGTGTGGGATGTTTCGATTGT	139
QY	127	TGCAGACAGCCCCCAAGCCCGCCCGCGCGACAGGGGAGAGGTGCAGAGAGCGACCCAGCC	186
Dp	140	TGTAGTAGCTGTTAGTTTGTCTGTCCGCGGTAGCGGAGAGGTGAGAGCGGATCGAGTT	199
QY	187	TCCCGGCGCACCCCGACCCCTTGCGCGACGCGCGCGACGACGAAGGGGCTGTTGCAGAACT	246
Dp	200	TTTTCGGTTATTCGATTTTATTTTGGGGGTAGCGCGCTAGAGAGTAAGGGGTTGGTGAATAT	259
QY	247	CGACCAACTCTACTCGCGCGCGCGCGCAGAGTGGGCTACCTGTCTACCGCGCGCGCAG	306
Dp	260	CGATTAATTTTATTTTCGCGCGCGGTAAAGTGGGTTATTTGTTTACGCGGCGGTGGAG	319
QY	307	GTTCTCTTGACACTCGAGCCAGATGTTGCGTGGGCACTTGCTGCTCTGTCGCCGACG	366
Dp	320	GTTTTTTTTGATTTGGAGCCAGATGGATTGCGTGGGATATGTTGGTTTCGTTGCTAGAG	379
QY	367	AGCGCGGACAGATCCGCGCTTGCGCCACCGGAGCGCACTGCTTATTCGGCGCACGTGGA	426
Dp	380	AGCGGAGACAGATCGTTTGGCTTATTCGAGATATGTTTATTCGGGGGTATAGGGA	439
QY	427	CGTAGTCCCGCTCTGCGCTGCTTTGACCTCTGTGGGGGTCTGCAGCGCTTCTTCC	486
Dp	440	CGTAGTTTTCGTTTTTGGTTTGTGTTTTGATTTTGTGGGGGTTTCCACGCTTTTTCG	499
QY	487	GGTCAGACAGCGCGGCTAACACCCAAAGCCACTCTCCGCGGACCTCGCGGAGAGGA	546
Dp	500	GGTTAAGTACCGCGGCTAATATTTAAAGTATTTCTTCGCGGATTTTGGGCGGAGAGGA	559
QY	547	AAAGGGCGCGCTGTACGGGAGTGGGTCCGACCGGATCTGCACGCTACACCCCGGAGG	606
Dp	560	AAAAGGGGCGCTGTACGGGAGTGGGTCTGTACGGAATTTGTAGGTTTATTTCCGCGAGG	619
QY	607	CTTCAGCTTGAAGGCGCTCGCGCGCGCGCGACAGTGTGAAGAACCCCGCGCTCACACGGA	666
Dp	620	TTTTAGTTTACGTTTTGTCTGTCGCGCGTTAGTTGCGAAATTTTCGCTTTATATTCGGA	679
QY	667	GGCCACGACGATCTCCGGCGCACAGCAACCCGAGCGGACGCGCACTGCGCTTCGA	726



DB 42 CGTATAGTGTATAGTATGATGATCGAATATATAT 1

RESULT 9  
ABO48542/C  
ID ABO48542 standard: DNA: 957 BP.  
XX  
AC ABO48542:  
XX  
DT 12-JUL-2002 (first entry)  
XX  
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 35133.  
XX  
XX Human: cytosine methylation: 5'-CpG-3': uracil: cytosine; diagnosis;  
KW drug: side effect: cancer; central nervous system; cardiovascular;  
KW gastrointestinal: respiratory system; single nucleotide polymorphism;  
KW SNP: cell differentiation; ds.  
XX  
XX Homo sapiens.  
XX  
XX WO200218632-A2.  
XX  
XX 07-MAR-2002.  
XX  
XX 01-SEP-2001: 2001WO-EP10074.  
XX  
XX 01-SEP-2000: 2000DE-1043826.  
XX  
XX 05-SEP-2000: 2000DE-1044543.  
XX  
XX (EPIC-) EPIGENOMICS AG.  
XX  
XX Olek A, Plepenbrock C, Berlin K, Guelty D;  
XX  
XX WPI: 2002-371829/40.  
XX  
XX Determining the degree of cytosine methylation in genomic DNA, useful  
PT for diagnosis and prognosis, comprises selective hybridization of  
PT amplicons from chemically treated DNA  
XX  
XX Claim 12: 56bp + sequence listing: 56bp; German.  
XX  
XX This invention describes a novel method for determining the degree of  
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
CC genomic sample of DNA. The sample is treated chemically to convert  
CC cytosine (C) but not methylated C, to uracil, then part of the genomic  
CC DNA that contains the target C is amplified to form a labeled amplicon.  
CC The amplicon is hybridised to two classes, each with at least one  
CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers  
CC and the degree of hybridisation to both classes is determined from the  
CC label on the amplicon. From the ratio of labels hybridised to the two  
CC classes of oligomers, the degree of methylation is calculated. The method  
CC is used: (i) for diagnosis and/or prognosis of side effects of  
CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders  
CC of the central nervous, cardiovascular, gastrointestinal and respiratory  
CC systems etc., particularly by detecting mutations or single nucleotide  
CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue  
CC types and for investigating cell differentiation. The method allows the  
CC methylation status of many C residues to be determined simultaneously.  
CC ABO13410-ABO54121 represent genomic DNA sequences used to illustrate the  
CC method for determining the degree of cytosine methylation described in  
CC the disclosure of the invention.  
XX  
XX Sequence 957 BP: 147 A: 102 C: 345 G: 363 T: 0 other;

Query Match 34.1% Score 517.2: DB 24: Length 957;  
Best Local Similarity 75.5% Pred. No. 3.1e-103;  
Matches 642: Conservative 0; Mismatches 208; Indels 0; Gaps 0;

OY 8 CAGCCACATATCTGCTGGTGGGCGTCCCTGCTGGCGTTCGCGCTGCCCG 67  
II IIIIIII III III IIIIIII III IIIIIII III IIIIIII III  
DB 850 CAAAGCAGTACTACTGCAATTAAGTCTCTACTACTAGCGGTTCGCGCTACCCCA 791  
XX  
OS 68 GCGGGGCTGCGCCCGCGGACACCTGCCAGATAAAGCGGCGACGCTCCGACTGCT 127  
XX

DB 790 ACCGGATGACCCCGCGCGACACTACCAAAATTAACGAAACACTCGACTACT 731  
II IIIIIII III IIIIIII III III III III IIIIIII III  
OY 128 GCAGACGCGCGCCAGCCCGCGGCGAGGAGAGAGAGAGAGAGAGAGAGAGAG 187  
II III IIIIIII IIIIIII III III III III III IIIIIII III  
DB 730 ACACACACCGCGCCAGCCCGCGGACGACAAAAAATAATACAAAAAGAGAGAGAG 671  
II IIIIIII IIIIIII III III III III III IIIIIII III  
OY 188 CCGGCGACCGCGCCAGCCCGCGGCGAGCGCGCAGAGAGAGAGAGAGAGAGAGAG 247  
II IIIIIII IIIIIII III III III III III IIIIIII III  
DB 670 CCGGACACCGCGCCAGCCCGCGGACGAGCGCAAAAAAATAATACAAAAATC 611  
II IIIIIII IIIIIII III III III III III IIIIIII III  
OY 248 GACCAACTTACTCCGCGCGCGGCAAGTGCGCTACTCTGTACGCGCGCGCGCAG 307  
II IIIIIII IIIIIII III III III III III IIIIIII III  
DB 610 GACCAACTTACTCCGCGCGCGGCAAGTGCGCTACTCTGTACGCGCGCGCGCA 551  
II IIIIIII IIIIIII III III III III III IIIIIII III  
OY 308 TTCCTGTTGAGCTGAGGAGATGTTGGGAGATGTTGGGAGATGTTGGGAGATG 367  
II IIIIIII III III III III III III III III IIIIIII III  
DB 550 TTCCTGTTGAGCTGAGGAGATGTTGGGAGATGTTGGGAGATGTTGGGAGATG 491  
II IIIIIII III III III III III III III III IIIIIII III  
OY 368 GCGGGGACGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 427  
II III IIIIIII IIIIIII III III III III III IIIIIII III  
DB 490 AACGAAACGATAGCGCGCTAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 431  
II IIIIIII IIIIIII III III III III III IIIIIII III  
OY 428 GGTAGTCCCGCGCTCTGCTGCTCTTTGACCTCTGTGGGGTCTGAGCGCTTTTCCG 487  
II IIIIIII IIIIIII III III III III III IIIIIII III  
DB 430 GCTAATCCCGCGCTCTGCTGCTCTTTGACCTCTGTGGGGTCTGAGCGCTTTTCCG 371  
II IIIIIII IIIIIII III III III III III IIIIIII III  
OY 488 GTCAAGCAGCGCGCTGACCGCTTAAGCGCACTGTGCGCGCGCGCGCGCGAG 547  
II IIIIIII IIIIIII III III III III III IIIIIII III  
DB 370 ATCAAAACGCGCGCGCTGACCGCTTAAGCGCACTGTGCGCGCGCGCGCGAG 311  
II IIIIIII IIIIIII III III III III III IIIIIII III  
OY 548 AAGGGCGCGTACGCGGAGTGGTCCGCGCGCGCGCGCGCGCGCGCGCGCG 607  
II IIIIIII IIIIIII III III III III III IIIIIII III  
DB 310 AAAAAACGCGTATACGAAATTAATCCGCGCGCGCGCGCGCGCGCGCGCG 251  
II IIIIIII IIIIIII III III III III III IIIIIII III  
OY 608 TTCAGCTTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 667  
II IIIIIII IIIIIII III III III III III IIIIIII III  
DB 250 TTCAGCTTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 191  
II IIIIIII IIIIIII III III III III III IIIIIII III  
OY 668 GCCGACGATGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 727  
II IIIIIII IIIIIII III III III III III IIIIIII III  
DB 190 ACCGACGATGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 131  
II IIIIIII IIIIIII III III III III III IIIIIII III  
OY 728 CTCCTGAGACGATCGCGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 787  
II IIIIIII IIIIIII III III III III III IIIIIII III  
DB 130 CTCCTGAGACGATCGCGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 71  
II IIIIIII IIIIIII III III III III III IIIIIII III  
OY 788 CCG 847  
II IIIIIII IIIIIII III III III III III IIIIIII III  
DB 70 CGACGACGCGCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 11  
II IIIIIII IIIIIII III III III III III IIIIIII III  
OY 848 TCCATGCGCG 857  
II IIIII III  
DB 10 TCCATACCG 1

RESULT 10  
ABO48543  
ID ABO48543 standard: DNA: 957 BP.  
XX  
AC ABO48543:  
XX  
DT 12-JUL-2002 (first entry)  
XX  
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 35134.  
XX  
XX Human: cytosine methylation: 5'-CpG-3': uracil: cytosine; diagnosis;  
KW drug: side effect: cancer; central nervous system; cardiovascular;  
KW gastrointestinal: respiratory system; single nucleotide polymorphism;  
KW SNP: cell differentiation; ds.  
XX  
XX Homo sapiens.  
XX

XX	MO200218632-A2.
XX	
PD	07-MAR-2002.
XX	
PF	01-SEP-2001; 2001WO-EPI0074.
XX	
PR	01-SEP-2000; 2000DE-1043826.
PR	05-SEP-2000; 2000DE-1044543.
XX	
PA	(EPID-) EPIGENOMICS AG.
XX	
P1	Olek A., Piepenbrock C., Berlin K., Guelzig D;
DR	WPI: 2002-371829/40.
XX	
PT	Determining the degree of cytosine methylation in genomic DNA, useful
PT	for diagnosis and prognosis, comprises selective hybridization of
PT	amplicons from chemically treated DNA -
PS	Claim 12: 56pp + Sequence Listing; 56pp: German.
XX	
CC	This invention describes a novel method for determining the degree of
CC	methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC	genomic sample of DNA. The sample is treated chemically to convert
CC	cytosine (C) but not methylated C, to uracil, then part of the genomic
CC	DNA that contains the target C is amplified to form a labeled amplicon.
CC	The amplicon is hybridized to two classes, each with at least one
CC	member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
CC	and the degree of hybridisation to both classes is determined from the
CC	C label on the amplicon. From the ratio of labels hybridised to the two
CC	classes of oligomers, the degree of methylation is calculated. The method
CC	is used: (i) for diagnosis and/or prognosis of side effects of
CC	therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
CC	of the central nervous, cardiovascular, gastrointestinal and respiratory
CC	systems etc., particularly by detecting mutations or single nucleotide
CC	polymorphisms (SNPs); and (ii) for differentiation of cell or tissue
CC	types and for investigating cell differentiation. The method allows the
CC	methylation status of many C residues to be determined simultaneously.
CC	AB013410-AB034121 represent genomic DNA sequences used to illustrate the
CC	method for determining the degree of cytosine methylation described in
CC	the disclosure of the invention.
XX	
S0	Sequence 957 BP; 363 A; 345 C; 102 G; 147 T; 0 other;
	Query Match            34.1%; Score 517.2; DB 24; Length 957;
	Best Local Similarity    75.5%; Pred. No. 3.1e-103;
	Matches 642; Conservative 0; Mismatches 208; Indels 0; Gaps 0
OY	8 CAGCGCACTATCTGCTGCTGGGTGGGGCTGCCGTGCTGTGGCGCTTCCGCCTCCGCTG 67
DB	108 CAACGACTATCTACTCTGTAATAAATGAATTCCTCACTACTATACGCGTTCCGCTCAAC 167
OY	68 GCCGGGCTGGCCCCCGCCGCGCACACTGCCAGAGATAAAGCCGGGCGAGCTCCGACTGCT 127
DB	168 ACCGGGATGACCCTCCGCGCGACACTTACCCAATAATAAACGACAACACTTCGACTACT 227
OY	128 GCAGCAGCGCCGACCCCGCCGCGCGCGAGGAGAGAGAGAGTGCAGAGCGAGCGGCTT 187
DB	228 ACAACACAGCGCCCAACCCCGCGCAGCAGCAAAAAAAAATAATACAAAAAGCAACCGAAC 287
OY	188 CCCGGCCACCGCCACCCCTTGCGCGCAGCGGCGCGAGCAGCAAGGGCTGTGTCAGAACATC 247
DB	288 CGCGACACCCCGACCCCTTAACGCAACGACCAAAAAAATACTAATTACAAAAATC 347
OY	248 GACCAACTTACTCCGGCGGGGCAAGGTGGGCTACCTGTACGCGGGGCGGCCGACAG 307
DB	348 GACCAACTTACTCCGAGCAGCAGCAAAAAAAACTACTCTGTCTACGCGAAGCAGCAAAA 407
OY	308 TTTCCTTTGGACCTGGAGCAGATGGTTGGTGGGCAATTCGATTCGCTGCCGCGAGA 367
DB	408 TTCCTTTAAACCTAAAGCAATATATTGTGATTAACATTAACACTTTCGTACCCGCAAAA 467
OY	368 GGCGGAGCAGTGGCGCCTTGCGCGCACCGAGCCACTGTTCTANTCGGGGACACATGGAC 427

Db	468	AACGAAAGGATACGCCCCCTAAGCGACCGAAACCACTACTTCTATGAAACACATTAAC	527
Qy	428	GGTAGTCCCGCTCTGCGCTGTTTGTAGCTCTGAGGGGCTTCGAGGGCTTCTTGCG	487
Db	528	GCTAATCCCCGCTCTGTAAGTATCTTTAACTCTATTAATAATCTCGACGACTTCTTGCG	587
Qy	488	GTCGAAGCAGCGGGCTTACACCTTAAGCACACTGCTGCGGGAGCCCTGCGGGAAGAA	547
Db	588	ATCAAAACACGCGCCCTACACCTTAACCACTACTACGCAACCCCTTAAGAAAAA	647
Qy	548	AAGGGGCGCGTGTACGGGGATGAGTCCGACAGGATCCTGACGCTACACCGCGAGGG	607
Db	648	AAAAAAGCGCTATACGAAATAATTAATCCGACGAATCTTACAGCTTACACCGGAAAC	707
Qy	608	TTCAAGCTTGAAGCCCTGCGCCGCGGCGAGCTGCCAAACCCCGCGTCCACACCGAG	667
Db	708	TTCAACCTTCGAAACCCCTACACCGCGCGGCGCACTACGAAACCCCGGTCACACCGAA	767
Qy	668	GCCACAGCAGCATGCTCCGGCGCACAGACACCCGAGCGGAGCGAGCAGTGGCTCGAG	727
Db	768	ACCCAGCAACATCTCCGACGACCAACACCCGAAACGACGACACTTAACCTGCGAA	827
Qy	728	CTCTGGAGCACTCCGCTCTGCGCCGCTGGGGCGTCCAGACCGACGAGCTGCGCG	787
Db	828	CTCTTAACCAATCCGCTCTTTCGCGCTTAACCACTCAAAACCGCAACGTAATAACGA	887
Qy	788	CGGCGCGCCGCTCCATCTCCCGGGCCCGCAGGTGAGCTGCTTCTGCTGAGCGC	847
Db	888	CGAGCAGCGCCGCTCATCTCCGGAACCGCAATAAATACTTTTAATACTAACGCG	947
Qy	848	TCCATGGCGC	857
Db	948	TCCATTAACGC	957
RESULT 11			
ABQ48544	ID	ABQ48544 standard; DNA; 957 BP.	
ABQ48544:			
XX	12-JUL-2002	(first entry)	
XX	XX		
DE	XX	Oligonucleotide for detecting cytosine methylation SEQ ID NO 35135.	
DE	XX		
XX	XX	Human: cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;	
KM	XX	drug; side effect; cancer; central nervous system; cardiovascular;	
KM	XX	gastrointestinal; respiratory system; single nucleotide polymorphism;	
KM	XX	SNP; cell differentiation; ds.	
OS	XX	Homo sapiens.	
PN	XX	MO200218632-A2.	
PD	XX	07-MAR-2002.	
PE	XX	01-SEP-2001; 2001MO-EPI0074.	
PR	XX	01-SEP-2000; 2000DE-1043826.	
PR	XX	05-SEP-2000; 2000DE-1044543.	
PA	XX	(EPiG-) EPIGENOMICS AG.	
PI	XX	Olek A. Piepenbrock C, Berlin K, Guetig D;	
DR	XX	WPI; 2002-371829/40.	
PT	XX	Determining the degree of cytosine methylation in genomic DNA, useful	
PT	XX	for diagnosis and prognosis, comprises selective hybridization of	
PS	XX	amplicons from chemically treated DNA	
PS	XX	Claim 12; 56pp + Sequence Listing; 56pp; German.	

XX This invention describes a novel method for determining the degree of  
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
 CC genomic sample of DNA. The sample is treated chemically to convert  
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic  
 CC DNA that contains the target C is amplified to form a labeled amplicon.  
 CC The amplicon is hybridised to two classes, each with at least one  
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers  
 CC and the degree of hybridisation to both classes is determined from the  
 CC label on the amplicon. From the ratio of labels hybridised to the two  
 CC classes of oligomers, the degree of methylation is calculated. The method  
 CC is used: (i) for diagnosis and/or prognosis of side effects of  
 CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders  
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory  
 CC systems etc., particularly by detecting mutations or single nucleotide  
 CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue  
 CC types and for investigating cell differentiation. The method allows the  
 CC methylation status of many C residues to be determined simultaneously.  
 CC AB013410-AB054121 represent genomic DNA sequences used to illustrate the  
 CC method for determining the degree of cytosine methylation described in  
 CC the disclosure of the invention.

XX Sequence 957 BP; 146 A; 102 C; 319 G; 390 T; 0 other;

Query Match 33.5%; Score 507.6; DB 24; Length 957;

Best Local Similarity 74.8%; Pred. No. 3.8e-101;

Matches 636; Conservative 0; Mismatches 214; Indels 0; Gaps 0;

7 GCAGCCACTATCTCTGCTGCGGTGGCGTCCCTGCTGCTGCGCTCCGCTGCCCT 66  
 107 GTAGCGTATATCTTCTTCTGCGGTGGCGTCTTCTTCTGCGCTTCTGCTTCTT 166  
 67 GCGCGGCTGCGCGCGCGCGCGACACCTCCAGATAAAGCCGGGACGCTTCGACTGC 126  
 167 GGTGCGGCTGCGGTGCTGCGCATATTTGTAGATAAAGTGGGTAGTTTTCATGTC 226  
 127 TCGACAGCGCGCGCGCGCGCGCGCGAGGAGGAGGTGACAGAGCGAGCGCGCC 186  
 227 TGTAGTAGTGTGTTATTTCTGCGCGGTAGGAGGAGGTGTAGACCGACTCGACT 286  
 187 TCCGCGCGACCG 246  
 287 TTTTCGTTATTTCTTATTTTGGCGGTAGCGCGCTAGAGATAGAGGTTGTGTAATAT 346  
 247 CGACCACTCTACTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 306  
 347 CGATTATTTATTTTGGCGCGCGGTAGGAGGCGGTATTTCTGTTACGCGCGCGTGG 406  
 307 GTTCCCTTGGAGCTGAGCGAGATGCTTCCGCTGCGCATTTGCGCTTCCCGCGCG 366  
 407 GTTTTGTGATTTTGGAGCGAGATGCTTCCGCTGCGGTATTTGCTTCTCTAGG 466  
 367 AGCGGCGAGAGTGCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 426  
 467 AGCGGCGAGAGTGCCTTGGCGGTATTCGAGATTTCTTTTATCGGCGGTATTCGA 526  
 427 CGGTAGTCCCGCTCTGCGCTCTTTAGACTCTGCGGCTTCGACGCGCTTCTTCC 486  
 527 CGTTAGTCTTCTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 586  
 487 GGTCAAGCG 546  
 587 GGTAAAGTGGCGCGGTATTTTAAAGTATTTTGGCGGTATTTTGGCGGTATTTG 606  
 547 AAAGGCGCGGTGAGGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 606  
 647 AAAGGCGCGGTGAGGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 706  
 607 GTTACGTTTGAAGCG 666  
 707 TTTTATTTTGAAGTGTGCGCGCGCGGTAGTTGCGAAATTTTGGCGGTATTCGGA 766  
 667 GCGCCGAGGATGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 726

DB 767 GCTTACGAGTATGTTTCCGCGGTATGATTAATTCAGACCGCGCGGTAGTATTCGTA 826  
 QY 727 GCCTTGACACATCGCTCTCCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 786  
 DB 827 GTTTTGGATGATGCTTCTTCTGCTGCTGCGGTAGCATCTTAAACGCGCGCG 886  
 QY 787 GCG 846  
 DB 887 GCGCGCGCGCGCGGTATTTTTCGCGGTGTTAGGTGAGGATTTTGTGTGTTACCC 946  
 QY 847 GTTCATGCGC 856  
 DB 947 GTTATGCGC 956

RESULT 12  
 AB048545/C  
 ID AB048545 standard; DNA; 957 BP.  
 XX AB048545;  
 AC 12-JUL-2002 (first entry)  
 XX  
 DT  
 XX  
 DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 35136.  
 XX  
 KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;  
 KW drug; side effect; cancer; central nervous system; cardiovascular;  
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;  
 KW SNP; cell differentiation; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200218632-A2.  
 XX  
 PD 07-MAR-2002.  
 XX  
 PF 01-SEP-2001; 2001WO-EP10074.  
 XX  
 PR 01-SEP-2000; 2000DE-1043826.  
 XX  
 PR 05-SEP-2000; 2000DE-1044543.  
 XX  
 PA (EPIG-) EPIGENOMICS AG.  
 XX  
 PI Olek A, Piepenbrock C, Berlin K, Guelig D;  
 XX  
 DR WPI: 2002-371829/40.  
 XX  
 PT Determining the degree of cytosine methylation in genomic DNA, useful  
 PT for diagnosis and prognosis, comprises selective hybridization of  
 PT amplicons from chemically treated DNA  
 XX  
 PS Claim 12; 56pp + Sequence Listing; 56pp; German.  
 XX  
 CC This invention describes a novel method for determining the degree of  
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
 CC genomic sample of DNA. The sample is treated chemically to convert  
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic  
 CC DNA that contains the target C is amplified to form a labeled amplicon.  
 CC The amplicon is hybridised to two classes, each with at least one  
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers  
 CC and the degree of hybridisation to both classes is determined from the  
 CC label on the amplicon. From the ratio of labels hybridised to the two  
 CC classes of oligomers, the degree of methylation is calculated. The method  
 CC is used: (i) for diagnosis and/or prognosis of side effects of  
 CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders  
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory  
 CC systems etc., particularly by detecting mutations or single nucleotide  
 CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue  
 CC types and for investigating cell differentiation. The method allows the  
 CC methylation status of many C residues to be determined simultaneously.  
 CC AB013410-AB054121 represent genomic DNA sequences used to illustrate the  
 CC method for determining the degree of cytosine methylation described in









```
Db 1049 GTGATACCTTTGGGATGGCTGATGTTGGAACGTGTGTGATCCGACAGAACCTGCTCCG 1108
QY 1203 TGATTGAAGCGATGGCGCTCCAGCGCTTCTGCTGCGTCAAGAAATCGGACATTTAC 1262
Db 1109 TCATTAGAAGATGATGGTTTACAGCTGCTTCACACAGGCCCATGAATTAGGCCACGTGT 1168
QY 1263 TTGGCCTCTCCCATGACGATTCACAATTCTGTGAAAGAGACCTTTGGTTCACAGAGATA 1322
Db 1169 TTAACTATGCCACATGATGATGCAAAAGCAGTGTCCAGCCTTAATGTTGAACCAAGATT 1228
QY 1323 AGCGCTTAATGTCTTCCATCCTTACCAGCATTTGATGCAATCTAAAGCCCTGGTCCAAATGCA 1382
Db 1229 CCCACATGATGGGCTCAATGCTTCCAACTGGACCCACAGCCGCTGTCTCTTGCA 1288
QY 1383 CTTGAGCGACATCACAGAATTCTGGATGATGGCCATGGTAACTGTTGCTGGACCTAC 1442
Db 1289 GTGCTTACATGATTAATCATTTCTTGATTAATGTTGATGGGAATGTTGATGGACAAGC 1348
QY 1443 CACGAAGCAGATCTG 1459
Db 1349 CTCGAATCCCATACAG 1365
```

Search completed: April 25, 2003, 16:50:49  
Job time : 675.931 secs



TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
2 (bases 1 to 5533)  
10438522  
J. Biol. Chem. 274 (33), 23443-23450 (1999)  
Abbaszade, I., Liu, R.-Q., Yang, F., Rosenfeld, S. A., Ross, O. H.,  
Link, J. R., Ellis, D. M., Tortorella, M. D., Pratta, M. A., Hollis, J. M.,  
Wynn, R., Duke, J. L., George, H. J., Hillman Jr., M. C., Murphy, K.,  
Barbata, H., Waisell, B. H., Copeland, R. A., Decicco, C. P., Bruchner, R.,  
Nagase, H., Ito, Y., Newton, R. C., Magolda, R. L., Trzaskos, J. M.,  
Hollis, G. F., Arner, E. C. and Burn, T. C.  
Direct Submission  
Submitted (10-Apr-1999) Applied Biotechnology, Dupont  
Pharmaceuticals Co., Experimental Station E336/237B, Wilmington, DE  
19880, USA  
FEATURES  
source  
1..5533  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
1..5533  
/gene="ADAMTS11"  
123..2915  
/gene="ADAMTS11"  
/note="ADAMTS11: ADAMTS family aggrecanase"  
/codon\_start=1  
/product="aggrecanase-2"  
/protein\_18="A4P49577.1"  
/db\_xref="GI:5733438"  
/translation="MLGMSALLCLAFPLPLAAGVPAATPAQDKAGQPPPTAAAAQPR  
RROGEVQAEAPPQHPHPLAQRKRSKGLVONIDLYSGGKRVGLVYAGGRPLDL  
ERDSGVJAGFAPAGGTSAPMRHSKFLVYREGSPFLPRACETPILVPA  
ARYTLPLRGPMAREEKGRVYGDGSAFLVYREGSPFLPRACETPILVPA  
HEHAPHSNPSGRALASOLLDSALSPAGSGGQTPRRRRRSISRROYELLIA  
ASMARLYRGLOHILTLASTANRLYSASLENIHLAVVYVYVGLDKLEYSKNA  
ATTLKNEFKWQHNLGDDEHYDAILFTREDLGHSCDTLGMADVLTCSPE  
SCAVLEDGLAAFTVAHEIGHLGLSHDSKFCERTGSTEDKRLMSILTSIDASK  
PMSKTSATLIEFLDGHNCILDLPRKOILGPELPCQYDADQOCMLTGPPEYSVC  
PGMDYCARLMCAVVGOCMCLTKLPAVEGPCGKGRICLOGKCDVTKRKYSTSS  
HGNSMSGMSGCSGSGGVOPARHCRNPNRGVYCKGRALVYSCSLMCPNP  
CKSPRHECEKAKNGYQSDAKVKTFFEVPRYAGVLPADYCKLTKRACTGYVYFSP  
KVTDTGECRPSNSVYCRKACVKTGCDIIGSKLOYDYGAGGGSNCTIVGFENK  
IDNGTAVNYSGWSHRDPLFHMGYSATKEILYQILATDPTKPLDVAVSFEVRSST  
PKVNSVTSHGSKVSGHSHTSOPQWVTGFWLACSRICDTHMTHTVQCODGNRKLAKGCP  
LSORPSATKCOLLKRC"  
BASE COUNT 1504 a 1270 c 1276 g 1483 t  
ORIGIN  
Query Match 95.8%; Score 1453; DB 9; Length 5533;  
Best Local Similarity 99.7%; Pred. No. 1,6e-251;  
Matches 1456; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 7 GCAGCCACTATGCTGCTGGGTGGGCGTCCCTGCTGCTGCTGCGCTTCCGCTGCGCCT 66  
DB 113 GCAGCCACTATGCTGCTGGGTGGGCGTCCCTGCTGCTGCTGCGCTTCCGCTGCGCCT 172  
QY 67 GGCAGGCGTGGGCGGCGGCGGCGACCTGCGAGATTAAGCGGGCAGCTCCGCACTGC 126  
DB 173 GGCAGGCGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 232  
QY 127 TGCAGCAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 186  
DB 233 TGCAGCAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 292  
QY 187 TCCCGGCGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 246  
DB 293 TCCCGGCGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 352  
QY 247 GCAGCAACTACTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 306  
|||||

DB 353 GCAGCAACTACTACTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 412  
QY 307 GTTCCCTTTGAGACTGAGAGCGAGATGTTGCGTGGGCACTTCTGCTGCGCCGAGG 366  
DB 413 GTTCCCTTTGAGACTGAGAGCGAGATGTTGCGTGGGCACTTCTGCTGCGCCGAGG 472  
QY 367 AGCGGGAGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 426  
DB 473 AGCGGGAGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 532  
QY 427 CGGTATGTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 486  
DB 533 CGGTATGTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 592  
QY 487 GGTCAAGCAGCGCGCGCTACACCTTAAGCCACTGCTGCGCGGAGCCCTGGCGGAGAGA 546  
DB 593 GGTCAAGCAGCGCGCGCTACACCTTAAGCCACTGCTGCGCGGAGCCCTGGCGGAGAGA 652  
QY 547 AAAGGGCGCGGTGACGCGGAGATGCTCCGACGAGATCTGACAGCTTACACCGCGAGGG 606  
DB 653 AAAGGGCGCGGTGACGCGGAGATGCTCCGACGAGATCTGACAGCTTACACCGCGAGGG 712  
QY 607 CTTCAAGCTTGAGAGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 666  
DB 713 CTTCAAGCTTGAGAGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 772  
QY 667 GGCACAGAGCATGCTCCGCGCGCACAGCAACCCGAGCGGACGAGCAGCTGGCTCGCA 726  
DB 773 GGCACAGAGCATGCTCCGCGCGCACAGCAACCCGAGCGGACGAGCAGCTGGCTCGCA 832  
QY 727 GCTCTTGGACAGTGGCTCTCTCTGCGCGCGCTGGGGCTGAGAACCGGACAGTGGTGGG 786  
DB 833 GCTCTTGGACAGTGGCTCTCTCTGCGCGCGCTGGGGCTGAGAACCGGACAGTGGTGGG 892  
QY 787 GCGGGGCGCGCGCTCATGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 846  
DB 893 GCGGGGCGCGCGCTCATGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 952  
QY 847 GTTCATGGCGCGGTGTATAGCGCGGCGCTGAGCATTAACCTGCTGACCGCTGGCTCCAT 906  
DB 953 GTTCATGGCGCGGTGTATAGCGCGGCGCTGAGCATTAACCTGCTGACCGCTGGCTCCAT 1012  
QY 907 GCGCAATAGGCTGTACAGCATGCTAGCATGAGAACCAACCATCCGCTGGCGGTGTGA 966  
DB 1013 GCGCAATAGGCTGTACAGCATGCTAGCATGAGAACCAACCATCCGCTGGCGGTGTGA 1072  
QY 967 GGTGGTGGCTAGGCGAGCAAGCAAGAGCTGAGAGTGAGCAAGAACGCTGGCCACAC 1026  
DB 1073 GGTGGTGGCTAGGCGAGCAAGCAAGAGCTGAGAGTGAGCAAGAACGCTGGCCACAC 1132  
QY 1027 ACTCAAGAACTTTGCAAGTGGCGAGCACCAACAGCAGCTGGAGATGACCATGAGGA 1086  
DB 1133 ACTCAAGAACTTTGCAAGTGGCGAGCACCAACAGCAGCTGGAGATGACCATGAGGA 1192  
QY 1087 GCACATAGATGAGCATATCTCTTACTGGGAGATTAATTTGGGCATTAATGTA 1146  
DB 1193 GCACATAGATGAGCATATCTCTTACTGGGAGATTAATTTGGGCATTAATGTA 1252  
QY 1147 CACCTTGGGAATGGCAGAGAGTGGAGCATATGTTCTCCAGAGCCAGCTGTGCTGAT 1206  
DB 1253 CACCTTGGGAATGGCAGAGAGTGGAGCATATGTTCTCCAGAGCCAGCTGTGCTGAT 1312  
QY 1207 TGAAGAGATGGCTTCACAGCAGCTTCACTGGCTCGACGAATGAGCATTTACTTGG 1266  
DB 1313 TGAAGAGATGGCTTCACAGCAGCTTCACTGGCTCGACGAATGAGCATTTACTTGG 1372  
QY 1267 CCTTCCCATGACGATTTCCAAATTTCTGTGAAGAGACCTTTGGTCCACGAAAGATAAGG 1326  
DB 1373 CCTTCCCATGACGATTTCCAAATTTCTGTGAAGAGACCTTTGGTCCACGAAAGATAAGG 1432  
QY 1327 CTTAATGTTTCAATCCTTACAGCATGATGATTAAGCCCTGGTCCAAATGACATTC 1386  
DB 1433 CTTAATGTTTCAATCCTTACAGCATGATGATTAAGCCCTGGTCCAAATGACATTC 1492  
|||||



QY 1267 CCTCTCCATGACGATTCCTCAATTCCTGTGAAGAGACCTTTGGTTCACAGAGATTAACG 1326  
|||||  
Db 1979 CCTCTCCATGACGATTCCTCAATTCCTGTGAAGAGACCTTTGGTTCACAGAGATTAACG 2038  
QY 1337 CTTAATGCTTCATCTCTTACCAGATTGATGATCTTAAGCCCTGGTCCAAATGACATTC 1386  
|||||  
Db 2039 CTTAATGCTTCATCTCTTACCAGATTGATGATCTTAAGCCCTGGTCCAAATGACATTC 2098  
QY 1387 AGCCGACATCAGAAATTCCTGGATGATGGCCATGATGTAAGTGTTCGTCGACCTACACG 1446  
|||||  
Db 2099 AGCCGACATCAGAAATTCCTGGATGATGGCCATGATGTAAGTGTTCGTCGACCTACACG 2158  
QY 1447 AAAGCAGATCTGGGGGGA 1467  
|||||  
Db 2159 AAAGCAGATCTGGGGGGA 2179  
RESULT 3  
AP001601/c 172558 bp DNA linear PRI 29-APR-2000  
LOCUS Homo sapiens genomic DNA, chromosome 21, clone:KB1346F10,  
DEFINITION AP001601.1 GI:7670555  
ACCESSION AP001601.1  
VERSION AP001601.1  
KEYWORDS Homo sapiens pre-pro-B cell cell\_line:FLB14-14 DNA, clone\_id:Keio  
SOURCE BAC library clone:KB1346F10.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 172558)  
AUTHORS Shimizu,N., Kudoh,Y. and Shlbuya,K.  
TITLE Homo sapiens genomic DNA, chromosome 21, clone:KB1346F10,  
JOURNAL AP001601.1  
REFERENCE 2 (bases 1 to 172558)  
AUTHORS Shimizu,N., Kudoh,Y. and Shlbuya,K.  
TITLE Direct Submission  
JOURNAL Submitted (04-APR-2000) Nobuyoshi Shimizu, Keio University, School  
of Medicine, Molecular Biology; 35 Shinanomachi, Shinjuku-Ku, Tokyo  
160-8582, Japan (E-mail:nshimizu@med.keio.ac.jp,  
Tel:81-3-3351-2370, Fax:81-3-3351-2370)  
FEATURES  
source  
1..172558  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/clone="KB1346F10"  
/cell\_line="FLB14-14"  
/cell\_type="pre-pro-B cell"  
/clone\_id="Keio BAC library"  
BASE COUNT 53953 a 32300 c 32498 g 53807 t  
ORIGIN  
Query Match 73.4%: Score 1113.4; DB 9; Length 172558;  
Best Local Similarity 99.9%: Pred. No. 1.3e-190;  
Matches 1114: Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 7 GCAGGCACTATGCTGCTCGGGTGGGCGTCCCTGCTGCTGTCGCGGCTCCGCTCCCT 66  
|||||  
Db 30657 GCAGGCACTATGCTGCTCGGGTGGGCGTCCCTGCTGCTGTCGCGGCTCCGCTCCCT 30598  
QY 67 GCGCGCGTGGGCGCGCGGACACTGCGGAGATTAAGCGGGGCGGCTCCGACTGC 126  
|||||  
Db 30597 GCGCGCGTGGGCGCGCGGACACTGCGGAGATTAAGCGGGGCGGCTCCGACTGC 30538  
QY 127 TGCAGCAGCCCGCCAGCCCGCGCGGCGGAGGAGAGAGTGCAGAGCGAGCCAGCC 186  
|||||  
Db 30537 TGCAGCAGCCCGCCAGCCCGCGCGGCGGAGGAGAGTGCAGAGCGAGCCAGCC 30478  
QY 187 TCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 246  
|||||  
Db 30477 TCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 30418

QY 247 CGACCACTTACTCCGGCGGCGCAAGGTGGGTACCTCTTACGGGGGCGGCGGCGGCGG 306  
|||||  
Db 30417 CGACCACTTACTCCGGCGGCGGCGCAAGGTGGGTACCTCTTACGGGGGCGGCGGCGGCGG 30358  
QY 307 GTTCCTCTTGGACCTGAGGAGATGTTGGTGGGCGGCTTCTGCTGCTGCTGCTGCTG 366  
|||||  
Db 30357 GTTCCTCTTGGACCTGAGGAGATGTTGGTGGGCGGCTTCTGCTGCTGCTGCTGCTG 30298  
QY 367 AGGGGGGAGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 426  
|||||  
Db 30297 AGGGGGGAGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 30238  
QY 427 CGGTAGTCCCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 486  
|||||  
Db 30237 CGGTAGTCCCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 30178  
QY 487 GGTCAAGCAGCGGCGGCTACACCTTAAGCCACTGCTGGCGGCGGCGGCGGCGGCGG 546  
|||||  
Db 30177 GGTCAAGCAGCGGCGGCTACACCTTAAGCCACTGCTGGCGGCGGCGGCGGCGGCGG 30118  
QY 547 AAAGGGCGGCGGTAAGGAGGATGGGTCGCGACGATCTGCGACGCTACACCGCGAGG 606  
|||||  
Db 30117 AAAGGGCGGCGGTAAGGAGGATGGGTCGCGACGATCTGCGACGCTACACCGCGAGG 30058  
QY 607 CTTCAGCTTCGAGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 666  
|||||  
Db 30057 CTTCAGCTTCGAGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 29998  
QY 667 GCGCCAGCAGCATGCTCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 726  
|||||  
Db 29997 GCGCCAGCAGCATGCTCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 29938  
QY 727 GCTCTTGAGCAGTCCGCTCTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 786  
|||||  
Db 29937 GCTCTTGAGCAGTCCGCTCTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 29878  
QY 787 GCGGCGGCGGCGGCTTCATCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 846  
|||||  
Db 29877 GCGGCGGCGGCGGCTTCATCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 29818  
QY 847 GTCCATGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 906  
|||||  
Db 29817 GTCCATGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 29758  
QY 907 GCGCAATAGCTGTACACCATGCTAGCATGAGAACCATCCGCTGCGGCTGGTGA 966  
|||||  
Db 29757 GCGCAATAGCTGTACACCATGCTAGCATGAGAACCATCCGCTGCGGCTGGTGA 29698  
QY 967 GGTGTGCTGCTAGGCGGCAAGAGACAGACCTGGAAGTGAAGAAAGCTGCCACAC 1026  
|||||  
Db 29697 GGTGTGCTGCTAGGCGGCAAGAGACAGACCTGGAAGTGAAGAAAGCTGCCACAC 29638  
QY 1027 ACTCAAGAACTTTTGCAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1086  
|||||  
Db 29637 ACTCAAGAACTTTTGCAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 29578  
QY 1087 GCACATGATGAGCTATCTGTTACTCGGAGG 1121  
|||||  
Db 29577 GCACATGATGAGCTATCTGTTACTCGGAGG 29543

RESULT 4  
AP001698/c 34000 bp DNA linear PRI 10-MAY-2000  
LOCUS AP001698 Homo sapiens genomic DNA, chromosome 21q, section 42/105.  
DEFINITION AP001698 AL163243 BA000005  
ACCESSION AP001698.1 GI:7768708  
VERSION AP001698.1  
KEYWORDS Homo sapiens DNA.  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

[illegible]

```
/note="L2"
/rpl_family="LINE/L2"
/rpl_type-DISPERSED
repeat_region
8515..8613
/note="(TACA)n"
/rpl_family="Simple_repeat"
/rpl_type-TANDEM
8614..8663
/note="(TACA)n"
/rpl_family="Simple_repeat"
/rpl_type-TANDEM
complement(9819..10245)
/note="L2"
/rpl_family="LINE/L2"
/rpl_type-DISPERSED
repeat_region
10905..10976
/note="L2"
/rpl_family="LINE/L2"
/rpl_type-DISPERSED
repeat_region
11003..11128
/note="L2"
/rpl_family="LINE/L2"
/rpl_type-DISPERSED
repeat_region
11601..11933
/note="LTR/L2"
/rpl_family="LTR/MaLR"
/rpl_type-DISPERSED
complement(13009..13204)
/note="LTR50"
/rpl_family="LTR"
/rpl_type-DISPERSED
repeat_region
14826..14871
/note="(A)n"
/rpl_family="Simple_repeat"
/rpl_type-TANDEM
complement(16472..17697)
/gene="ADAMTS5"
/number=1
misc_feature
16743..17389
/note="Cpg Island"
complement(17576..17697)
/gene="ADAMTS5"
18711..19114
/note="Cpg Island"
complement(21277..21404)
repeat_region

Query Match      73.4%: Score 1113.4; DB 9; Length 340000;
Best Local Similarity 99.9%: Pred. No. 1.2e-190;
Matches 1114; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 367 AGCGGAGACAGTGGCCCTGGCCACCGGAGACCACTGCTTCTATCGGGGACAGTGA 426
|||||
DB 17225 AGCGGAGACAGTGGCCCTGGCCACCGGAGACCACTGCTTCTATCGGGGACAGTGA 17166
QY 427 CGTGAATCCCGCTCTCTGCTGCTTTGACCTCTGTGGGGGTCTGACGGCTTCTCCG 486
|||||
DB 17165 CGTGAATCCCGCTCTCTGCTGCTTTGACCTCTGTGGGGGTCTGACGGCTTCTCCG 17106
QY 487 GGTCAAGCAGCGGGGCTACACCTTAAGCCACTGCTGGCGGAGACCTGGGGAGAGA 546
|||||
DB 17105 GGTCAAGCAGCGGGGCTACACCTTAAGCCACTGCTGGCGGAGACCTGGGGAGAGA 17046
QY 547 AAAGGGGCGCGTGTACCGGGATGGGTCCGACAGATCTCTGACACCCCGCGAGGG 606
|||||
DB 17045 AAAGGGGCGCGTGTACCGGGATGGGTCCGACAGATCTCTGACACCCCGCGAGGG 16986
QY 607 CTTCAAGCTTGAGAGCCCTGCGCCCGCGCGGCGCAGCTGCGAACCCTCCACACCGGA 666
|||||
DB 16985 CTTCAAGCTTGAGAGCCCTGCGCCCGCGCGGCGCAGCTGCGAACCCTCCACACCGGA 16926
QY 667 GGGCCAGCAGCATGGCTCGGGGCGACAGCAACCGGAGCGGAGCGACGTGGCTCGCA 726
|||||
DB 16925 GGGCCAGCAGCATGGCTCGGGGCGACAGCAACCGGAGCGGAGCGACGTGGCTCGCA 16866
QY 727 GCTCTTGACACAGTCGCTCTCTGCTGCTGGGGGCTTGAGACCGACAGCTGTGGCG 786
|||||
DB 16865 GCTCTTGACACAGTCGCTCTCTGCTGCTGGGGGCTTGAGACCGACAGCTGTGGCG 16806
QY 787 GGGGGGGGGCGCTGCATCTCCGGGGCGGCGGCGGCGGAGGAGCTGCTGTGGCTGAGCC 846
|||||
DB 16805 GGGGGGGGGCGCTGCATCTCCGGGGCGGCGGCGGAGGAGCTGCTGTGGCTGAGCC 16746
QY 847 GTCAATGGCGCGGTGTATGGCGGGGCGCTGCAGCATTAACCTGTGACCTGGCTTCAT 906
|||||
DB 16745 GTCAATGGCGCGGTGTATGGCGGGGCGCTGCAGCATTAACCTGTGACCTGGCTTCAT 16686
QY 907 CGCCATTAAGCTGTACAGCGCATGCTAGCATGAGAACCAATCCCGCTGGCTGTGA 966
|||||
DB 16685 CGCCATTAAGCTGTACAGCGCATGCTAGCATGAGAACCAATCCCGCTGGCTGTGA 16626
QY 967 GGTGGGTGGTGTAGGAGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1026
|||||
DB 16625 GGTGGGTGGTGTAGGAGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 16566
QY 1027 ACTCAAGAACTTTTGCACATGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1086
|||||
DB 16565 ACTCAAGAACTTTTGCACATGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 16506
QY 1087 GCACATGATGACGATATCTGTTTACTGCGAGG 1121
|||||
DB 16505 GCACATGATGACGATATCTGTTTACTGCGAGG 16471

RESULT 5
AC025891/c 163811 bp DNA linear HTG 10-SEP-2000
LOCUS Homo sapiens chromosome 21 clone RP11-588J22 map 21, WORKING DRAFT
DEFINITION SEQUENCE, 18 unordered pieces.
ACCESSION AC025891
VERSION AC025891.3 GI:10047872
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 163811)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 21, clone RP11-588J22
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 163811)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
```



TITLE  
 JOURNAL  
 COMMENT

Submitted (16-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Sep 10, 2000 this sequence version replaced gi:7342190.  
 All repeats were identified using RepeatMasker:  
 Smt, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

-----  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Genome Center  
 Center code: MIBR  
 Web site: <http://www-seq.wi.mit.edu>  
 Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
 -----  
 Project Information  
 Center project name: L8251  
 Center clone name: 588.J.22  
 -----

----- Summary Statistics -----  
 Sequencing vector: M13: M77815: 100% of reads  
 Chemistry: Dye-terminator Big Dye: 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 150751 bases at least Q40  
 Consensus quality: 157173 bases at least Q30  
 Consensus quality: 159998 bases at least Q20  
 Insert size: 163000; agarose-fp  
 Insert size: 162111; sum-of-ctrls  
 Quality coverage: 4.5 in Q20 bases; agarose-fp  
 Quality coverage: 4.5 in Q20 bases; sum-of-ctrls

-----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 18 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

\* 1  
 \* 27637: contig of 27637 bp in length  
 \* 27638 27737: gap of 100 bp  
 \* 27738 28822: contig of 1085 bp in length  
 \* 28823 28922: gap of 100 bp  
 \* 28923 32108: contig of 3186 bp in length  
 \* 32109 32208: gap of 100 bp  
 \* 32209 34033: contig of 1825 bp in length  
 \* 34034 34133: gap of 100 bp  
 \* 34134 39035: contig of 4902 bp in length  
 \* 39036 39135: gap of 100 bp  
 \* 39136 43900: contig of 4765 bp in length  
 \* 43901 44000: gap of 100 bp  
 \* 44001 50223: contig of 6223 bp in length  
 \* 50224 50323: gap of 100 bp  
 \* 50324 55366: contig of 5033 bp in length  
 \* 55367 55466: gap of 100 bp  
 \* 55467 60972: contig of 5506 bp in length  
 \* 60973 61072: gap of 100 bp  
 \* 61073 68474: contig of 7402 bp in length

FEATURES	source	Location/Qualifiers
*	68475	68574 : gap of 100 bp
*	68575	75868 : contig of 7294 bp in length
*	75868	75968 : gap of 100 bp
*	75968	85343 : contig of 9375 bp in length
*	85344	85443 : gap of 100 bp
*	85444	98666 : contig of 13223 bp in length
*	98667	98766 : gap of 100 bp
*	98767	111217 : contig of 12451 bp in length
*	111218	111317 : gap of 100 bp
*	111318	128495 : contig of 11178 bp in length
*	128496	128595 : gap of 100 bp
*	128596	139817 : contig of 11222 bp in length
*	139818	139917 : gap of 100 bp
*	139918	156952 : contig of 17035 bp in length
*	156953	157052 : gap of 100 bp
*	157053	163811 : contig of 6759 bp in length.
Location/Qualifiers		
1. 163811		
/organism="Homo sapiens"		
/db_xref="taxon:9606"		
/chromosome="21"		
/map="21"		
/clone="RP11-568J22"		
/clone_lib="RPC1-11 Human Male BAC"		
1. 27637		
misc_feature	/note="assembly-fragment"	
clone_end:SP6		
vector_side:left"		
misc_feature	27738..28822	
	/note="assembly-fragment"	
misc_feature	28923..32108	
	/note="assembly-fragment"	
misc_feature	32209..34033	
	/note="assembly-fragment"	
misc_feature	34134..39035	
	/note="assembly-fragment"	
misc_feature	39136..43900	
	/note="assembly-fragment"	
misc_feature	44001..50223	
	/note="assembly-fragment"	
misc_feature	50324..55366	
	/note="assembly-fragment"	
misc_feature	55467..60972	
	/note="assembly-fragment"	
misc_feature	61073..68474	
	/note="assembly-fragment"	
misc_feature	68575..75868	
	/note="assembly-fragment"	
misc_feature	75969..85343	
	/note="assembly-fragment"	
misc_feature	85444..98666	
	/note="assembly-fragment"	
misc_feature	98767..111217	
	/note="assembly-fragment"	
misc_feature	111318..128495	
	/note="assembly-fragment"	
misc_feature	128596..139817	
	/note="assembly-fragment"	
misc_feature	139918..156952	
	/note="assembly-fragment"	
misc_feature	157053..163811	
	/note="assembly-fragment"	
clone_end:r7		
vector_side:right"		
BASE COUNT	50540 a 30683 c 30303 g 50580 t 1705 others	
ORIGIN		
Query Match	73.2%:	Score 1110.2; DB 2; length 163811;
Best Local Similarity	99.7%:	Pred. No. 4.7e-190;
Matches 11121: Conservative	0: Mismatches 3; Indels 0; Gaps 0;	



Db 68 GTCCTGCTGCTCCCTGCGCCGTGACAGCCCGCCGCGCACCCTGCCAGGATAAACACAG 127  
 QY 112 GCAGCCTCCAGCTGCTGACGACCGCCCGCCGCGCGGCGAGGGAGAGAGGTGA 171  
 Db 128 GCAAGCTTCAGGCTGACGACCGCCGCGAGCGAGCGCGAGGGAGAGAAACAG 187  
 QY 172 GGAGGAGCGGAGCCTCCCGGCCACCCCGCCCTGCGCGCAGCGCGAGAGAGAGG 231  
 Db 188 GGAGGAGG-----GGCATTATACACCTTGGCGGGGACGAGGAGCGGGG 235  
 QY 232 GCTGTGTCAGAAATCAGCAACTCTACTCCGCGCGCGGCAAGGTGGCTTACCTGCTTA 291  
 Db 236 GCTGTGTCATATATAGCAACTCTACTCTGCGTGCGCAAGGTGGCTTACCTGCTTA 295  
 QY 292 CGCGGCGCGCGCAGGCTCTGCTGACCTGAGGAGAGATGGTGGTGGGATGCTGCT- 350  
 Db 296 CGCGGCGCGCGGAGGTTCTGCTGACCTGAGAGAGATGACACAGTGGTGGCTGCTG 355  
 QY 351 --GCTTGTGCGCGAGAGAGCGGAGAGTCCGCTGCGCCACCGAGCCACTGCTT 408  
 Db 356 TAGCATCTGTTACTGACGAGAGAGGCTGAGCCATCTCTGGCCACCGGGTCACTGTT 415  
 QY 409 CATGTGGGCGACAGTGGAGAGGTAGTCCCGCTCTGCTGCTGTTGACCTGCTGGGG 468  
 Db 416 CTACAGAGGACCGCTGGAGCGGAGCCCTGATCCCTAGCTGCTTGGACCTGCGCGG 475  
 QY 469 TCTGACGCGCTTCTGCGCGTCAAGCAGCGGCGCTACACCTAAAGCCACTGTCGCGG 528  
 Db 476 TCTGATGCGCTTCTTGGAGTCAAGCATCGCGCTACACTTAAAGCCACTGCTGGTGG 535  
 QY 529 ACCCTGCGCGGAGAGAAAGAGGGCGCTGTACGGGGATGGGTCCGACGATCTCTGA 588  
 Db 536 GACCTGCGGAGAGTATGA-----CGAATTATGGGATGGTCTTCCCGACTCTGGA 589  
 QY 589 CCTTACACCGCGGAGGCTTACGCTTGAAGCCCTGCGCGCGCGCGGAGCTGCGCAAC 648  
 Db 590 TCTTACCAACCGCGAGGGCTTGAAGTTCAGGCGCCCTGCGCGCGGAGCTGCGCAAC 649  
 QY 649 CCGCGGCTCACAGCGGAGCGGAGCATGCTCCGCGCAGAGCAACCGGAGCGAGG 708  
 Db 650 TCTGTGATCCCATCTGCGGCGCCCAAGAGAGCCCTGCGTGCACAGTATGATGAGAGAG 709  
 QY 709 CGCAGCATGCGCTGCGACCTCTTGGACCAAGTCCGCTCTGCGCGCTGGGGGCTCAG 768  
 Db 710 CTCAGGCTGCGCGCGGACACTGCTGACCACTCAGCTTCTGCGCATCTGGGAAACGCGG 769  
 QY 769 ACCGAGATGCTGGCGCGGCGCGGCGGCTGATCCCGCGCGCGGAGTGGAGGT 828  
 Db 770 ACCCTGAGCTTGGTGGAGCGGAGGCGGCTTCATCTCCAGGGCGCGCAGGTGAGCT 829  
 QY 829 GCTTCTGTGCTGAGCGCTCCATGCGCGGCTTGTATGGCGCGGCGCTCAGCATTA 888  
 Db 830 CCTTGTGTGGGTGAGCTGCTGACGAGATGATATGGGGCGGCGCTCAGCATTA 889  
 QY 889 GCTGACCTTGGCTTCTCATGCGCAATAGCTGTACAGCCATGCTAGCATGAGAACCAT 948  
 Db 890 GCTGACCTTGGCTTCTCATGCGCAATAGCTGTACAGCCATGCTAGCATGAGAACCAT 949  
 QY 949 CGGCTGCGCGGCTGAGAGGTGGTGTAGGCGACAGAGCAAGAGCTGGAAGTGA 1008  
 Db 950 CGGCTGCGCGGCTGAGAGGTGGTGTAGGCGACAGAGCAAGAGCTGGAAGTGA 1009  
 QY 1009 CAGAAGCCTGCGCACCACTCAGAACTTTTGAAGTGGCGAGCAGACACACACAGCT 1068  
 Db 1010 CAGAATGCGGCGAGACCTCAAGAACTTTTGAAGTGGCGAGCAGACACACATTA 1069  
 QY 1069 GAGAGATGACATGAGAGAGCATAGCATAGCTATCTGTTTACTCGGAGAGATTATG 1128  
 Db 1070 AAGGAGTGTATCAGAGAGACATAGCATAGCTATCTGTTTACCCGAGAGATTATG 1129  
 QY 1129 TGGGATCATATGATGACACCTGGAGATGGAGAGTGGAGACATATGTTCTCAG 1188  
 Db 1130 TGGGATCATATGATGACACCTGGAGATGGAGAGTGGAGACATATGTTCTCAG 1189

QY 1189 GCGACAGTGTGCTGATGTAAGAGAGATGGCCCTCCACGAGCCTTCACTGTGGCTACGA 1248  
 Db 1190 GCGACAGTGTGATGATGTAAGAGATGGCCCTCCATGAGCCTTCACTGTGGCTACGA 1249  
 QY 1249 AATCGACATTTACTTGGCTCTCCATGAGATTCCTCAATTTCTGTGAAGAGACCTTTGG 1308  
 Db 1250 AATGGGATTTACTTGGCTCTTCTCATGACATTCCTCAATTTCTGTGAAGAGACCTTTGG 1309  
 QY 1309 TTCCACAGAAGATAGCGCTTAATGTCTTTCATCTTACCAGCATTTGATGATCAAGCC 1368  
 Db 1310 TACTACAGAAGACAGCGTTTATGTCTTTCATCTTACCAGCATTTGATGATCAAGCC 1369  
 QY 1369 CTGGTCAATGACATTTGACGCGCCATTCACAGATTCCTGGATGATGGCATGTAAGT 1428  
 Db 1370 CTGGTCAATGACATTTGACGCGCCATTCACAGATTTCTGGATGATGGCATGTAAGT 1429  
 QY 1429 TTTGCTGACCTACACAGAAAGCAGATCTGGG 1461  
 Db 1430 TTTGCTGACCTACACAGAAAGCAGATTTGGG 1462

RESULT 7  
 HSA328940 663 bp DNA linear PRI 18-JUL-2002  
 LOCUS  
 DEFINITION Homo sapiens genomic sequence surrounding Not1 site, clone  
 ACCESSION AJ328940  
 VERSION AJ328940.1 GI:15873358  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 1 (bases 1 to 663)  
 Kuleshko, A.S., Gizatullin, R.Z., Al-Amin, A.N., Wang, F., Khasha, S.M.,  
 Podowski, R.M., Matushkin, Y.G., Gyanchandani, A., Muravenko, O.V.,  
 Levitsky, V.G., Kolchanov, N.A., Protodopov, A.I., Khashuba, V.I.,  
 Kiselev, L.L., Masserman, W., Wahlstedt, C., and Zaborovsky, E.R.  
 Not1 flanking sequences: a tool for gene discovery and verification  
 of the human genome  
 Nucleic Acids Res. 30 (14), 3163-3170 (2002)

JOURNAL  
 PUBMED 12136098  
 REFERENCE 2 (bases 1 to 663)  
 Zaborovsky, E.R.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-MAY-2001) Microbiology and Tumorbiology Centre,  
 Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77,  
 Sweden

FEATURES  
 source 1..663  
 location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="NL1-AE24R"

BASE COUNT 96 a 230 c 223 g 113 t 1 others

Query Match 43.2% Score 655.6; DB 9; Length 663;  
 Best Local Similarity 99.2%; Pred. No. 4.7e-108;  
 Matches: 658; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 297 GCGGCGGAGGTTCTCTTGGACCTGGAGCAGATGTTGCGGGGCAATTTGCTGCTG 356  
 Db 1 GCGGCGGAGGTTCTCTTGGACCTGGAGCAGATGTTGCGGGGCAATTTGCTGCTG 60  
 QY 357 TGGCGCGAGAGGCGGAGAGTGGCGGCTGCGCGCCACCGAGACACTCTTATCGG 416  
 Db 61 TGGCGCGAGAGGCGGAGAGTGGCGGCTGCGCGCCACCGAGACACTCTTATCGG 120  
 QY 417 GCACAGTGAAGGTAGTCCCGGCTCTGCGTCTTGAACCTGTGGGGGCTCGAG 476  
 Db 121 GCACAGTGAAGGTAGTCCCGGCTCTGCGTCTTGAACCTGTGGGGGCTCGAG 180



```

* 9818 10990: contig of 1173 bp in length
* 10991 11090: gap of unknown length
* 11091 12374: contig of 1284 bp in length
* 12375 12474: gap of unknown length
* 12475 14015: contig of 1541 bp in length
* 14016 14115: gap of unknown length
* 14116 15184: contig of 1069 bp in length
* 15185 15284: gap of unknown length
* 15285 16352: contig of 1068 bp in length
* 16353 16452: gap of unknown length
* 16453 17565: contig of 1113 bp in length
* 17566 17665: gap of unknown length
* 17666 18881: contig of 1216 bp in length
* 18882 18981: gap of unknown length
* 20281 20281: contig of 1300 bp in length
* 20381 20381: gap of unknown length
* 20382 22077: contig of 1696 bp in length
* 22078 22177: gap of unknown length
* 22178 23683: contig of 1506 bp in length
* 23684 23783: gap of unknown length
* 23784 25795: contig of 2012 bp in length
* 25796 25895: gap of unknown length
* 25896 27372: contig of 1477 bp in length
* 27373 27472: gap of unknown length
* 27473 29722: contig of 2250 bp in length
* 29723 29822: gap of unknown length
* 29823 31021: contig of 1199 bp in length
* 31022 31121: gap of unknown length
* 31122 33076: contig of 1955 bp in length
* 33077 33176: gap of unknown length
* 33177 34212: contig of 1036 bp in length
* 34213 34312: gap of unknown length
* 34313 35921: contig of 1609 bp in length
* 35922 36021: gap of unknown length
* 36022 37182: contig of 1161 bp in length
* 37183 37282: gap of unknown length
* 37283 38411: contig of 1129 bp in length
* 38412 38511: gap of unknown length
* 38512 39662: contig of 1151 bp in length
* 39663 39762: gap of unknown length
* 39763 41322: contig of 1560 bp in length
* 41323 41422: gap of unknown length
* 41423 43597: contig of 2175 bp in length
* 43598 43697: gap of unknown length
* 43698 45968: contig of 2271 bp in length
* 45969 46068: gap of unknown length
* 46069 47774: contig of 1706 bp in length
* 47775 47874: gap of unknown length
* 47875 50411: contig of 2537 bp in length
* 50412 50511: gap of unknown length
* 50512 52298: contig of 1787 bp in length
* 52299 52398: gap of unknown length
* 52399 54500: contig of 2102 bp in length
* 54501 54600: gap of unknown length
* 54601 57079: contig of 2479 bp in length
* 57080 57179: gap of unknown length
* 57180 58935: contig of 1756 bp in length
* 58936 59035: gap of unknown length
* 59036 60855: contig of 1820 bp in length
* 60856 60955: gap of unknown length
* 60956 63043: contig of 2088 bp in length
* 63044 63143: gap of unknown length
* 63144 65472: contig of 2329 bp in length
* 65473 65572: gap of unknown length
* 65573 68209: contig of 2637 bp in length
* 68210 68309: gap of unknown length
* 68310 70346: contig of 2037 bp in length
* 70347 70446: gap of unknown length
* 70447 73105: contig of 2659 bp in length
* 73106 73205: gap of unknown length
* 73206 74857: contig of 1652 bp in length
* 74858 74957: gap of unknown length
* 74958 76751: contig of 1794 bp in length

```

```

* 76752 76851: gap of unknown length
* 76852 79730: contig of 2879 bp in length
* 79731 79830: gap of unknown length
* 79831 82718: contig of 2888 bp in length
* 82719 82818: gap of unknown length
* 82819 85120: contig of 2302 bp in length
* 85121 85220: gap of unknown length
* 85221 85568: contig of 2348 bp in length
* 85569 87668: gap of unknown length
* 87669 90655: contig of 2987 bp in length
* 90656 90755: gap of unknown length
* 90756 94335: contig of 3580 bp in length
* 94336 94435: gap of unknown length
* 94436 97832: contig of 3397 bp in length
* 97833 97932: gap of unknown length
* 97933 100711: contig of 2779 bp in length
* 100712 100811: gap of unknown length
* 100812 103501: contig of 2690 bp in length
* 103502 103601: gap of unknown length
* 103602 107061: contig of 3460 bp in length

```

Query Match 38.7% Score 587: DB 2: Length 184454:  
 Best Local Similarity 72.7% Pred. No. 5,4e-96;  
 Matches 712: Conservative 0: Mismatches 255: Indels 12: Gaps 2:

```

OY 143 CCCCAGGCGGCGGAGGAGAGGTGACAGAGAGCCAGCCGCGGCGCACCCGCGAC 202
|||
Db 136018 CCCATGANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 136077
OY 203 CCCCTGGCGCAGCGCGCGAGAGCAAGGGCGTGTGAGAAATGACCACTACTCC 262
|||
Db 136078 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 136137
OY 263 GGGGGGCGCAAGGGGGGCTACTGCTAGCGGGGCGCGAGGTCCTTGAGACTG 322
|||||
Db 136138 GGGGCTGCAAGGGGCTACTGCTAGCGGGGCGCGAGGTCCTTGAGACTG 136197
OY 323 GAGGAGATGTTGGTGGGATGCTGGCTCTGCGCCGCGAGAGCGGGAGAGTGGC 382
|||||
Db 136198 GAGAGGATGACAGAGTGGGTCGTGGTGGCATCTTACTGAGAGGCGTGAAGCA 136257
OY 383 CCTTGGCGCCACCGGAGCCACTGCTTATCGGGGCGACATGAGCGGTAGTCCCGCTCT 442
|||||
Db 136258 TCCCTGGCGCCACCGGAGCCACTGCTTACAGAGGCGCTGGAGCGCCCTCGATCC 136317
OY 443 CTGGCTGCTTTGACCTCTGTGGGGTCTGACGCGCTTCTGCGGTCAGACGACGCGCG 502
|||||
Db 136318 CTAGCTGCTTTGACCTCTGTGGGGTCTGATGCTTTCGACGTCAGCATGCGCGC 136377
OY 503 TACACCTTAAGCCACTGCTGCGCGGACCGTGGGAGGAGAAAGAGGCGCGGTATC 562
|||||
Db 136378 TACACTGTAAGCGCGCTTCTGCGGGTCTGCGGACAGTCCGAA-----CGAGTTTAC 136431
OY 563 GGGATGAGTCCGACGAGATCTGACGCTTACACCGCGAGGGCTTACGTTGAGAGCC 622
|||||
Db 136432 GGGATGAGTCCGACGAGATCTGACGCTTACACCGCGAGGGCTTACGTTGAGAGCC 136491
OY 623 CTGGCGCGCGCGCGGACCTGCGGAAACCCCGCGTCCACCGAGGGCGCGACGATGCT 682
|||||
Db 136492 CTGGCGCGCGCGCGGACCTGCGGAAACCCCGCGTCCCGCGTGGGCGCGAAGAGCC 136551
OY 683 CCGCGGACGACGACCGGAGGCGGACGACGACCTGCTGCGGCTTGGACGACGATCC 742
|||||
Db 136552 TCGGTGCACTACTTCTAGGCGACGACGACGACGACGACGACGACGACGACGAC 136611
OY 743 GCTCTCTGCGCGCGTGGGGGCTGAGACCGGACGAGCTGTGGCGGGCGCGCGCGCTCC 802
|||||
Db 136612 GCTTCTCTGCGCGCGTGGGGGCTGAGACCGGACGAGCTGTGGCGGGCGCGCGCGCTCC 136671
OY 803 ATCTCCGCGCGCGCGGAGGAGGAGCTGCTTGTGCTGACGCGGCTGAGCGCGGTTG 862
|||||
Db 136672 ATCTCCAGGCGCGCGGAGGAGGAGCTGCTTGTGCTGACGCTTCCATGCGCAAGATG 136731

```



OY	357	TCGCCGAGAGAGCCGGAGAGATGGCGCCCTGGCGCCACCGAGAGCCACATGCTTCATCGAG	416
Db	61	TCGCCGAGAGAGCCGGAGAGATGGCGCCCTGGCGCCACCGAGCCACATGCTTCATCGAG	120
OY	417	GCACAGTGAAGCGTATGTCGCCGCTCTGTGGCTGTCTTTGACCTCTGTGGGGGTCTCGACG	476
Db	121	GCACAGTGAAGCGTATGTCGCCGCTCTGTGGCTGTCTTTGACCTCTGTGGGGGTCTCGACG	180
OY	477	GCTTCTTGGCGGTCAAGCAGCGCGCCTACACCCCTAAAGCACACTGCTGGCGGAGACCTGAG	536
Db	181	GCTTCTTGGCGGTCAAGCAGCGCGCCTACACCCCTAAAGCACACTGCTGGCGGAGACCTGAG	240
OY	537	CGGAGGAGAAAAGGGCGCGCTGTACGGGGATGGTCCGCACGATCCTCGACGCTTACA	596
Db	241	CGGAGGAGAAAAGGGCGCGCTGTACGGGATGGGCGCCGACAGGATCCTGTACGATGCA	300
OY	597	CCCCGAGAGGCTTCAGCTTTCGAGGCGCTTCGCGCCCGCGCCACGATCGCGAAACCCCGGAT	656
Db	301	CCCCGAGAGGCTTCAGCTTTCGAGGCGCTTCGCGCCCGCGCCACGATCGCTATTGGGAAACCCCG	359
OY	657	CCACACCGGAGGCCACGACGATGCTCCG	685
Db	360	TTATCCGGAGGCTATATAAGGTGTCCG	388

```

RESULT 11
AF317415          809 bp      mRNA      linear      MAM 19-NOV-2000
LOCUS              AF317415
DEFINITION         Oryctolagus cuniculus aggrecanase-2 (ADAMTS-11) mRNA, partial cds.
ACCESSION          AF317415
VERSION            AF317415.1  GI:11225592
KEYWORDS
SOURCE
ORGANISM           Oryctolagus cuniculus.
                   Oryctolagus cuniculus
                   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
REFERENCE           1 (bases 1 to 809)
AUTHORS            Goad,D.L. and Goad,M.E.
TITLE              Molecular cloning of lapine aggrecanase-2 (ADAMTS-11) from
                   articular chondrocytes
JOURNAL            Unpublished
REFERENCE          2 (bases 1 to 809)
AUTHORS            Goad,D.L. and Goad,M.E.
TITLE              Direct Submission
JOURNAL            Submitted (27-OCT-2000) Genetics Institute, MS G3001, One Burt
                   Road, Andover, MA 01810, USA
FEATURES
SOURCE             1..809
                   Location/Qualifiers

```

```

gene
/ gene="ADAMTS-11"
<1..>809
<1..>809
/ gene="ADAMTS-11"
/ codon_start=3
/ product="aggrecanase-2"
/ protein_id="AAG33062.1"
/db_xref="GI:11225593"
/ translation="TLGADVGTICSPRSCAVIEDDGLHAFTVAHETGHLGLSHL
DSKCEENFGSTEDKRLMSILTSIDAKPWSKCTSATITFELDDGHGNCILDVPRKQ
ILGEBELFGQTYDATQOCLNLFPEEYVTSCHMDVCARLMGCVVRGQAVCLTKLPA
EGTGEGRICLOGKCVDKTKKKYYSSTSHNMGCARLMGCVVRGQAVCLTKLPA
NPAPRNSRYCTGKRAIYRSCSYTPCPANGKSFRIHGECEAKNGYQSDAKG"

```

	Query Match	Similarity	Best Local	Score	DB 4:	Length
	Matches	293, Conservative	0, Mismatches	22, Indels	0, Gaps	0, Gaps
QY	1146	ACACCCCTGGGAATGCGACAGCTTGGACCATATGTTCTCCAGAGCGCACTGTGCTGTCA	1205			
DB	1	ACACCCCTGGGAATGCGACAGCTTGGACCATATGTTCTCCAGAGCGCACTGTGCTGTCA	60			

QY	1206	TTGAAAGCATGGGCTCCACGACGCTTCACGTGCTCAGCAATTCGACATTTACTTG	1265
Db	61	TTGAAGATGATGGCTCCACGCGGCTTTCACCGTGGCTCAGCAATTTGACATCTCCTTG	120
QY	1286	GCCCTCTCCCATGACGATTCCAAAATTCGTGAAAGACCTTTGGTTCACAGAAATGATAGC	1325
Db	121	GCCCTCTCTCATGACGATTCCAAAATTCGTGAAAGAGAACTTTGGCTCCACAGAAACAAGC	180
QY	1326	GCTTAATGTCTTCACATCTTACACGAGATTTGATGATCTTAAGCGCTGGTCCAAATGACACTT	1385
Db	181	GCTTAATGTCTATCCATCTCTACACGATATGATGATGCTTCCAAGCCCTTGTTCCAGTGCACCT	240
QY	1386	CAGCCACCATCAGAAATTCCTGGATGATGGCCATGGTAACTGTTTGCTGACCTACAC	1445
Db	241	CAGCCACCATCAGGAATTCCTGGATGATGGCCATGAAACTGTTTGCTGATGTACAC	300
QY	1446	GAAGCAGATCTCGGGGGGGA	1467
Db	301	GAAGCAGATCTCGGGGGGGA	322

RESULT	12			
LOCUS	E58656			
DEFINITION	Novel metallic protease.	2853 bp	DNA	linear
ACCESSION	E58656			
VERSION	E58656.1	GI:18629878		
KEYWORDS	JP 2001017183-A/4.			
SOURCE	Homo sapiens.			
ORGANISM	Homo sapiens.			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
AUTHORS	1 (bases 1 to 2853)			
TITLE	Yamaai,N., Nishimura,K. and Sasamata,M.			
JOURNAL	Novel metallic protease			
COMMENT	Patent: JP 2001017183-A 4 23-JAN-2001;			
	YAMANOUCHI PHARMACEUT CO LTD			
	OS Homo sapiens (human)			

```

PF      09-JUL-1999 JP 1999J196584
PR
PI      NOBORU YAMAJI, KOICHI NISHIMURA, MIHO SASAMATA
PC      C12N15/09,C07K16/40,C12NI/15,C12NI/19,C12N5/10, PC
C12N9/50,C12Q1/37,
PC      C12N15/00,C12N5/00
CC
FH      Key      Location/Qualifiers
FT      source    1..2853
FT
FT      Location/Qualifiers
      1..2853
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
BASE COUNT      704 a      734 c      816 g      599 t
ORIGIN

```

Query Match	17.6%	Score 267	DB 6	Length 2853
Best Local Similarity	53.4%	Pred. No. 2.9e-38		
Matches	586	Conservative	0	Mismatches 305; Indels 6; Gaps 1;
QY	363	CAGAGGGGCGAGCATGCGCCCTGGCCACACCGAGCAGCTTATCGGGACAG	422	
Db	275	CCGGGTCGAGAGCGCGCTCCGGAACGACGACTGGCCACTTACTCGGACCG	334	
QY	423	TGGACGTAAGTCCCGCTCTGCGCTCTTGGACTCTGTGGGGGTCGACGCTTCT	482	
Db	335	TGAATGGCGATCCGACTCGGCTGCGCCGCTCAGCGCTGCGAGGGCGCTGCGCGCGCT	394	
QY	483	TGCGGGTCAAGACGCGCGCTACACCTTAAAGCCACTGCTGGCGGAGACCTTGGCGGAGG	542	
Db	395	TCTACTGCTGGGGGAGGCGTATTTCATCAGCCCTCTCCCGCCGACGAGCGCCCTCG	454	

```

OY 543 AAGAAAAAGGGGCGGTGTACGGGGATGGGTCCGACGGATCTCGACGTCTACACCCGCG 602
DB 455 CCACGGCCCCCAGGAGGAGAGCCGCGGACACACTACAGTTCCACCTTCGCGGGCGGA 514
OY 603 AAGGCTTCAGCTTCGAGGCGCTCCGCGCGCGGCGGACGCTCGCAAAACCCCGGCTCCAC 662
DB 515 ATCGGCGAGGCGCAGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 574
OY 663 CGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 722
DB 575 GGAAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 634
OY 723 CGCAGCTCTTGAGACAGTCCGCTCTCTCGCGCGCTGGGGGCTGAGAGACCGAGCGTGT 782
DB 635 GGTCCCGGAGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 688
OY 783 GGGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 842
DB 689 GCATAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 748
OY 843 AGCGCTCCATGCGCGGCGGCTGTATGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 902
DB 749 ACCAGCTCGATGGCAGAAATTCACAGGCGAGGTCTAAAGATTACCTTCTCAGCTGTCTT 808
OY 903 CCATGCGCCATAGGCTGTACAGCCATGCTAGCATGAGAGACCAATCCGCGCTGGCGTGG 962
DB 809 CCGTGGCAGCCAGATTTGTACAAACCCAGCATTTCTTAATTCAGTTACCTGCTGGTGG 868
OY 963 TGAAGGTGTGTGTGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1022
DB 869 TGAAGATCTTGTGTATCCAGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 928
OY 1023 CCACACTCAAGAACTTTTGCAGAGTGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGT 1082
DB 929 TCAGCTCTGGGAACTTTTGCAGAGTGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGT 988
OY 1083 AGGAGCATACAGTACAGTATCTCTGTTTACTCGGAGAGATTTATGTGGCATCTTAT 1142
DB 989 CAGACCACTATGACACAGCAATTTCTTTCACGAGAGAGAGAGAGAGAGAGAGAGAGAT 1048
OY 1143 GTGACACCCCTGGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1202
DB 1049 GTGATACCTCTTGGGATGGCTGATGTTGGAGCTGTGTGATCCAGAGAGAGAGAGAGAG 1108
OY 1203 TGATTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1262
DB 1109 TCATAGAGATGATGTTTACAAAGCTGCTTCCAGCAGAGAGAGAGAGAGAGAGAGAG 1168
OY 1263 TTGGGCTCTCCCATGACATTCGCAAAATTCGTGGAAGAGAGAGAGAGAGAGAGAGAG 1322
DB 1169 TTAACATGCGCATGATGATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1228
OY 1323 AGCGCTTATGCTTCCATTCCTTACAGCATTTGATGATCTTAAGCCCTGGTCCAAATGCA 1382
DB 1229 CCCACATGATGGCGGCTGATGCTTCCAACTGGAGCAGAGAGAGAGAGAGAGAGAGAG 1288
OY 1383 GTTACGAGCAGCATCAGAGATTCCTGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAT 1442
DB 1289 GTGCTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1348
OY 1443 CAGGAAAGAGATCTGT 1459
DB 1349 CTCAGAAATCCCATACAG 1365

```

```

RESULT 13
AF060152 3430 bp mRNA linear PRI 11-AUG-1999
LOCUS Homo sapiens METH1 protein (METH1) mRNA, complete cds.
DEFINITION AF060152
ACCESSION AF060152.1 GI:5725505
VERSION
KEYWORDS

```

```

SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 3430)
AUTHORS
Vazquez, F., Hastings, G., Ortega, M.A., Lane, T.F., Oikemus, S.,
Lombardo, M., and Iruela-Arispe, M.L.
TITLE
METH-1, a human ortholog of ADAMTS-1, and METH-2 are members of a
new family of proteins with angio-inhibitory activity
J. Biol. Chem. 274 (33), 23349-23357 (1999)
JOURNAL
99367466
MEDLINE
10438512
PUBMED
2 (bases 1 to 3430)
REFERENCE
Vazquez, F., Hastings, G., Ortega, M.-A., Lane, T. F., Lombardo, M.,
Oikemus, S., and Iruela-Arispe, M. L.
AUTHORS
Direct Submission
TITLE
Submitted (16-APR-1998) Pathology, Beth Israel Deaconess Medical
Center, 99 Brookline Avenue, Boston, MA 02214, USA
JOURNAL
Local/Qualifiers
FEATURES
1..3430
/codon_start=1
/product="METH1 protein"
/protein_id="AAD48080.1"
/db_xref="GI:5725506"
/translation="MGNAERAPGSRSPGPEVPTLLLAALAVSDALGPSEDEELV
VPELEAPAGHCTTTLRLHAPODLDLEPSPSLAPGFTLONTGRRSGSTPEPD
LAHCFYSGTVNGDPSAALSLGCVGAFLLLEAFYFIOLPASRLATPAAGERP
PAPVQPHILRNRRGDYGTGTYDDEPRPTGKAREDEDEBGTGEBEGPMSQDPA
LOGVQPTGTSIRKKRVSRRYVERMLVADQSMABEHGSLHYLLTLESVARLY
KHPISIRNSVSLVAVKILVITIDEOKGPEVTSNAALTLNFCNMORQHNPPSDRDAHYD
TAILFTYDQDLSQCDTLGMAVDGTVCDPDRSCSVIEDDLQAAFTTAHLGIVFPM
PHDAKOCASLNGVNDOSHMAAASMLNIDHSPSPASAVMTFTSLNDRGECMDRP
ONPIOLPGLDPTGTYDANROCFTEGSESKICPPAASCTSLMCTGSGVLYVOTRH
FPMADGTCGCGKACIKNGKYNKTRDRHFDPRFSGMCMGPMGDCFRGCGVQYTM
RECNPPPKNGKATCEGRVATRSCLNEDCDNNGKTFREOCENHNEFSKASVSGP
AVENIPYAGVSPDRCKLICOAGGIGFVLOKRVVDGTPCSPHDSVYCVQGCVA
CGCDITIDSKKKFKGCVGCGGSGTCKKISGVSATSKPGYHDIITIPGATNIYKORN
QRCGRNNGSFLAIKAADGTYILNGDYTLSTLEODIMYKGVLYRSGSALEIRSPS
PLKEPLTIOVLTVCNALRPKIKYTFYFKKKKESFNALPTFSAVYIEEGESKCEIG
WQRLVLCROINQGPASECAKEVAPSTRPCADHPRCOMOLGEMSSCKTGKGYKRR
SLKLSHDGVLHSRSCDPLKRPKHFIDFTMACS"
BASE COUNT 877 a 845 c 971 g 736 t 1 others
ORIGIN
Query Match 17.6%; Score 267; DB 9; Length 3430;
Best Local Similarity 53.4%; Pred. No. 2.9e-38;
Matches 586; Conservative 0; Mismatches 505; Indels 6; Gaps 1;

```

```

OY 353 CAGAGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 422
DB 512 CCGGCTCCGAGAGCGCGGCTCCGGAACGAGCAGCTGGCGCAGCTTCTACATCCGCGCAG 571
OY 423 TGGAGGTAAGTCCCGGCTCTGCTGTCTTTGACCTCTGTGGGGGCTTCGAGGCTTCT 482
DB 572 TGAATGGCGATCCGAGCTGGCTGGCGGCTGCTGGAGGCGGCTGGCGGCT 631
OY 483 TCCGCTCAAGAGAGCGGCTACACCTTAAGGCACTGCTGGCGGAGAGAGAGAGAGAGAG 542
DB 632 TCTACCTCTGGGGAGGAGGATTTTATTCACAGCGGCTGCGCGGCGGAGAGAGAGAGAG 691
OY 543 AAGAAAGAGGCGGCTTACCGGAGTGGTCCGAGAGATCCGAGCTTACACAGAGAGAGAG 602
DB 692 CCACCGCGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 751
OY 603 AAGGCTTACGCTTCAGAGGCGGCGGCGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAG 662

```



Db 752 ATCCGAGGCGGACGCTAGCGGACGCTGGGGGTGTGGAGCAGACGCGCGGCGACTG 811  
 QY 663 CGGAGGCGGACGAGATGTCGCGGCGACAGACACCGGCGGCGGCGGCGGCGCT 722  
 Db 812 GGAAGCGGAGACCGAGACGAGACGAGAGGAGCTGAGGCGAGAGCGAAGGCGCTCACT 871  
 QY 723 CGCAGCTTTGGACGAGTCCGCTCTCTCGCGCTGGGGGCTCAGAGACCGGAGAGCTGT 782  
 Db 872 GGTGCGCGGAGACCGGCGGAGCTGCAAGGCGGTAGAGACCGCCACAGGA-----ACTGGAA 925  
 QY 783 GCGGCGGCGGCGGCGGCTCCATCTCCCGGCGCGGCGGAGGTGAGCTCTTGTGGCTG 842  
 Db 926 GCATAGAAAGAAAGGAGATTGTGTCCAGTCAACCGCTATGTGSAACACATGCTTGTGGCAG 985  
 QY 843 AGCGTCCATGGCGGCGGCTGTATGGCGGCGGCGGCTGAGCATTAAGCTGAGCTGCGCT 902  
 Db 986 ACCAGTCGATGGCAGAAATTCACAGGAGTGTCTAAGCATTAAGCTTCTCAGGCTTGT 1045  
 QY 903 CCATGCGCAATAGGCTGTACAGCAGTGTAGCATGAGACACATCCGCGTGGCGGCTG 962  
 Db 1046 CGGTGGCAGCGGAGATTGTACAAACACCCAGCATTCGTATTAATCAATGAGCTGTGGTGG 1105  
 QY 963 TGAAGGTGCTGTGTAGCGGCGACAGAGACAGGCGCTGAGAGCAAGACGCTGCCA 1022  
 Db 1106 TGAAGATCTTGTGTATCCAGATGAAAGAGAGGCGGCGGAGTGAACCTCATGCTGCC 1165  
 QY 1023 CCACACTCAAGAACTTTGCAAGTGGGAGCAGACACACAGCAGCTGGGAGATGACATG 1082  
 Db 1166 TCACTCTGGGAGCTTTGCAACTGGCAGAGCAGACACACCCAGCTGACCGGAGT 1225  
 QY 1083 AGGAGCACTAGCATGAGCTATCTCTGTACTCGGGAGATTTATGTGGCATTCAT 1142  
 Db 1226 CAGAGCACTATGACAGACAGAAATTTCTTTTACAGACAGACTTGTGTGGTCCAGACAT 1285  
 QY 1143 GTGACACCTGGGAATGCGAGAGCTTGGGACCATATGTTCTCAGAGCGGAGCTGTCTG 1202  
 Db 1286 GTGATACCTTTGGGATGGGTGATGTGTGGAAGTGTGTATCCGAGACAAACCTGCTCG 1345  
 QY 1203 TGAATGAAGAGCTGGGCGCTCCAGGAGCTTCACTGTGGCTACAGAAATCGGAGATTC 1262  
 Db 1346 TCAATGAAATATATGTTTACAAAGCTGCTTCACACACCCCATGATTAAGGCGAGT 1405  
 QY 1263 TTGGGCTTCCCATGAGCATTCCTCAAAATTCGTGAAGAGACCTTTGGTTCACAGAGATA 1322  
 Db 1406 TTAACATGCAATATGATGATGAAGACAGTGTGCCACTTATGTGTGAACAGAGATT 1465  
 QY 1323 AGCGCTTAATGTCTTCATCTTACAGCATGTGATGATCTAAGCCCTGGTCCAAATGCA 1382  
 Db 1466 CCCAGATGATGGCTCAATGCTTTCCAACTGGACACACAGCCGCTTGTCTCTTGA 1525  
 QY 1383 CTTGAGCCCACTCAAGATCTCTGATGATGGCAGCTGATGAGCTGTGGGAGACTGAC 1442  
 Db 1526 GTGCTCACTGATTTACATCATTTCTGATTAATGATGATGAGGAGATGATGAGACAG 1585  
 QY 1443 CACGAAGCAGATCTG 1459  
 Db 1586 CTCAGATCCCATACAG 1602

## RESULT 14

LOCUS E29406 4014 bp DNA linear PAT 18-JUN-2001  
 DEFINITION Novel integrin ligand ITGL-TSP.  
 ACCESSION E29406  
 VERSION E29406.1 GI:13025553  
 KEYWORDS JP 1999018784-A/1.  
 SOURCE unidentified.  
 ORGANISM unidentified.  
 REFERENCE 1 (bases 1 to 4014)  
 AUTHORS Zudekka,L.J., Steven,H.T., James,A.F., Greg,A.H. and Jonathan,A.T.  
 TITLE Novel integrin ligand ITGL-TSP

## JOURNAL

Patent: JP 1999018784-A 1 26-JAN-1999;  
 SMITHKLINE BEECHAM CORP, SMITHKLINE BEECHAM P L C, UMAN GENOME SCI

## COMMENT

OS Unidentified  
 PN JP 1999018784-A/1  
 PD 26-JAN-1999  
 PR 15-APR-1998 JP 1998104400  
 PR 24-APR-1997 US 08/845496  
 PI ZUDEKKA L, JONAKKU, STEVEN H TORURI, JAMES A FOMMORUDO, PI GREG  
 A HEISHTINGUSU,  
 PI JONATHAN A TERRET  
 PC C12N15/09, A61K39/395, A61K45/00, A61K45/00, A61K45/00,  
 PC A61K45/00,  
 PC A61K45/00, A61K48/00, C07K16/18, C12N1/21, C12P21/02, PC  
 C1201/02, C07K14/47,  
 PC C07K14/745, C07K14/78, (C12N1/21, C12R1/19), (C12P21/02, C12R1/19),  
 PC C12N15/00  
 CC Strandedness: Single;  
 CC Topology: linear;  
 CC Key  
 FH source  
 FT source

## FEATURES

source Location/Qualifiers  
 1..4014  
 /organism="unidentified"  
 /db\_xref="taxon:32644"

BASE COUNT 1007 a 1006 c 1135 g 866 t  
 ORIGIN

Query Match 17.6%; Score 267; DB 6; Length 4014;  
 Best Local Similarity 53.4%; Pred. No. 2.8e-38;

Matches 586; Conservative 0; Mismatches 505; Indels 6; Gaps 1;

QY 363 CAGGAGCGGCGAGAGTGGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 422  
 Db 791 CCGGGTCCGAGACGCGCTTCCGAAACGACCTGGCGGAGCTTCTACCTCGGCGGCGG 850  
 QY 423 TGGAGCGTGTGCGGCGCTGTGCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 482  
 Db 851 TGAATGCGCATTCGAGCTGGGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 910  
 QY 483 TCGCGGTCAAGCAGCGCGGCGCTACACCTAAAGCACTGCTGCGGCGGCGGCGGCGG 542  
 Db 911 TCTACTGTGTGGGAGGCGGTATTCATCCAGCCGCTGCGGCGGCGGCGGCGGCGGCTG 970  
 QY 543 AAGAAAGGCGGCGGCTGTACGGGATGGTCCGACAGATCTGACAGCTTACACCGCG 602  
 Db 971 CCACGCGCGCGCGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCA 1030  
 QY 603 AGGCGTTCAGCTTTCAGGCGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 662  
 Db 1031 ATGCGGAGGCGGAGCTAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1090  
 QY 663 CGGAGGCGGCGGAGAGATGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 722  
 Db 1091 GGAAGCGGAGACCGAAGACGAGAGCAAGAGGAGATGAGGCGGAGAGCGAAGGCGCTCACT 1150  
 QY 723 CGCAGCTTTGGACCAAGTCCGCTCTCTGCGGCTGGGGGCTCAGAGCGGAGAGCTGT 782  
 Db 1151 GGTGCGCGGAGGAGCGGCGGAGCTGCAAGGCGGTAGGAGCGGCGGCGGCGGCGG 1204  
 QY 783 GCGGCGGCGGCGGCGGCGGCTCATCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 842  
 Db 1205 GCATAGAAAGAAAGGAGATTGTGTCCAGTCAACCGCTATGTGSAACACATGCTTGTGGCAG 1264  
 QY 843 AGCGGTTCATGGCGGCGGCTGTATGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGCT 902  
 Db 1265 ACCAGTCGATGGCAGAAATTCACAGGAGAGTGTCTAAGCATTAAGCTTCTCAGCTGTTT 1324  
 QY 903 CCATGCGCAATAGGCTGTACAGCAGCATGAGTGAAGACACATCCGCGCTGGCGGCTGG 962  
 Db 1325 CGGTGCGACCGAGATTGTACAAACACCCAGCATTCGTATTAAGCTTGTGCGGTGG 1384





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 25, 2003, 21:42:46 ; Search time 91.2303 Seconds  
(without alignments)  
4568.371 Million cell updates/sec

Title: US-10-050-200-1  
Perfect score: 1359  
Sequence: 1 gaattcgcatgtccacagac.....atgacaaagggtagtcac 1359

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA.\*

1: /cgn2.6/ptodata/2/lna/5A\_COMB.seq:\*  
2: /cgn2.6/ptodata/2/lna/5B\_COMB.seq:\*  
3: /cgn2.6/ptodata/2/lna/6A\_COMB.seq:\*  
4: /cgn2.6/ptodata/2/lna/6B\_COMB.seq:\*  
5: /cgn2.6/ptodata/2/lna/PTCUS\_COMB.seq:\*  
6: /cgn2.6/ptodata/2/lna/Backfill1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1318.4	97.0	4192	US-09-122-126B-1	Sequence 1, Appli
2	280.2	20.6	3706	US-09-484-970B-58	Sequence 58, Appli
3	280.2	20.6	4676	US-09-130-491-1	Sequence 1, Appli
4	270.8	19.9	4858	US-09-392-184-1	Sequence 1, Appli
5	246.4	18.1	3638	US-09-369-364A-8	Sequence 8, Appli
6	233.8	17.2	3126	US-09-392-184-7	Sequence 7, Appli
7	228.4	16.8	739	US-09-369-364A-10	Sequence 10, Appli
8	180.6	13.3	3002	US-09-369-364A-1	Sequence 1, Appli
9	175.8	12.9	3250	US-09-122-126B-14	Sequence 14, Appli
10	137.4	10.1	5357	US-09-392-184-5	Sequence 5, Appli
11	134.4	9.9	3218	US-09-369-364A-6	Sequence 6, Appli
12	132.4	9.7	703	US-09-392-184-6	Sequence 6, Appli
13	124	9.1	2625	US-09-369-364A-14	Sequence 14, Appli
14	111	8.2	5804	US-09-369-364A-12	Sequence 12, Appli
15	62.8	4.6	385	US-09-392-184-23	Sequence 23, Appli
16	56.8	4.2	2848	US-09-369-364A-4	Sequence 4, Appli
17	52	3.8	2023	US-09-491-522-6	Sequence 6, Appli
18	52	3.8	2450	US-09-491-522-2	Sequence 2, Appli
19	52	3.8	6692	US-09-491-522-1	Sequence 1, Appli
20	47	3.5	3885	US-09-369-364A-16	Sequence 16, Appli
21	44.8	3.3	3489	US-08-728-323A-1	Sequence 1, Appli
22	44.8	3.3	3489	US-08-298-568-1	Sequence 1, Appli
23	44.8	3.3	32207	US-08-770-379-20	Sequence 20, Appli
24	44.8	3.3	32207	US-08-757-669A-20	Sequence 20, Appli
25	44.8	3.3	32207	US-09-230-371A-20	Sequence 20, Appli
26	44	3.2	1520	US-09-369-364A-3	Sequence 3, Appli
27	42.6	3.1	502	US-09-392-184-12	Sequence 12, Appli

28	41.6	3.1	2114	US-09-130-491-7	Sequence 7, Appli
29	41.4	3.0	1476	US-09-930-872-1	Sequence 1, Appli
30	41.4	3.0	3675	US-09-930-872-3	Sequence 3, Appli
31	41.4	3.0	4042	US-09-930-872-5	Sequence 5, Appli
32	41	3.0	2594	US-08-989-385-2	Sequence 2, Appli
33	40.2	3.0	7812	US-09-368-590-1	Sequence 1, Appli
34	39.8	2.9	733	US-09-392-184-15	Sequence 15, Appli
35	39.8	2.9	3691	US-09-211-704A-3	Sequence 3, Appli
36	39.6	2.9	543	5273901-6	Patent No. 5273901
37	39.4	2.9	1462	US-09-434-288-4	Sequence 4, Appli
38	39.4	2.9	2522	US-09-038-389A-1	Sequence 1, Appli
39	39.4	2.9	2522	US-09-611-781-1	Sequence 1, Appli
40	39.2	2.9	1926	US-09-249-585A-2	Sequence 2, Appli
41	39.2	2.9	2580	US-09-050-863-2	Sequence 2, Appli
42	39.2	2.9	2580	US-09-359-081-2	Sequence 2, Appli
43	39.2	2.9	5452	US-09-130-114-1	Sequence 1, Appli
44	39.2	2.9	9600	US-08-910-647-1	Sequence 1, Appli
45	39.2	2.9	9600	US-09-620-925-1	Sequence 1, Appli

## ALIGNMENTS

RESULT 1									
US-09-122-126B-1									
; Sequence 1, Application US/09122126B									
; Patent No. 6451575									
; GENERAL INFORMATION:									
; APPLICANT: Bristol-Myers Squibb Company									
; TITLE OF INVENTION: AGGRECAN DEGRADING METALLO PROTEASES									
; FILE REFERENCE: DM6909									
; CURRENT APPLICATION NUMBER: US/09/122,126B									
; CURRENT FILING DATE: 1998-07-24									
; NUMBER OF SEQ ID NOS: 21									
; SOFTWARE: PatentIn version 3.0									
; SEQ ID NO 1									
; LENGTH: 4192									
; TYPE: DNA									
; ORGANISM: Homo sapiens									
; FEATURE:									
; NAME/KEY: CDS									
; LOCATION: (406)..(2916)									
US-09-122-126B-1									
Query Match									
Best Local Similarity 97.0%; Score 1318.4; DB 4; Length 4192;									
Matches 1331; Conservative 0; Mismatches 21; Indels 0; Gaps 0;									
OY	7	GCATGTCACAGACAGCGTCGATCCCGGAGGAGGCTTGCAAGGCGCTGCTGGGA	66						
DB	403	GCATGTCACAGACAGCGTCGATCCCGGAGGAGGCTTGCAAGGCGCTGCTGGGA	462						
OY	67	GCCAACCTGCTCTGCTGCTCCCATTTGCGCTCTCTGCTGCTGCTGCTGCTG	126						
DB	463	GCCAACCTGCTCTGCTGCTCCCATTTGCGCTCTCTGCTGCTGCTGCTGCTG	522						
OY	127	CTACCTGCTGCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	186						
DB	523	CTACCTGCTGCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	582						
OY	187	GAGATGCTGCTTCCAGAGAGCTCAAGGAGCGTCTGCTGCTGCTGCTGCTGCTG	246						
DB	583	GAGATGCTGCTTCCAGAGAGCTCAAGGAGCGTCTGCTGCTGCTGCTGCTGCTG	642						
OY	247	AGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	306						
DB	643	AGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	702						
OY	307	TCCGATGTCAGAGTCAAGGAGGCTGACAGTCACTGCTGCTGCTGCTGCTGCTG	366						
DB	703	TCCGATGTCAGAGTCAAGGAGGCTGACAGTCACTGCTGCTGCTGCTGCTGCTG	762						
OY	367	GGTGAAGCAGAGCCTGCACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	426						

```

Db 763 GGTGGAGCAGACCTGGCAGCTACCTGAGCTGACCATGAGATCCGAGTCGGTG 822
Oy 427 GCATCTCTGACCTGGGATGGGGAGCCCTGTAGGGCTTACATATATGGGGGGCTGAA 486
Db 823 GCATCTCTGACCTGGGATGGGGAGCCCTGTAGGGCTTACATATATGGGGGGCTGAA 882
Oy 487 CTCACCTTCACAGCCCTGGAGGAGGACCCCTTAATCTGCTGGGGGACCTGGGCTCAC 546
Db 883 CTCACCTTCACAGCCCTGGAGGAGGACCCCTTAATCTGCTGGGGGACCTGGGCTCAC 942
Oy 547 ATCTACCGCCGGAAGAGTCTCTGACCGCTCAAGTCCCATGTGCAACGCTCAAGCTCCT 606
Db 943 ATCTACCGCCGGAAGAGTCTCTGACCGCTCAAGTCCCATGTGCAACGCTCAAGCTCCT 1002
Oy 607 CTTGGAAGCCCGAGCCCGAGACCCGGAAGAGCCGCTTGGCTTCACTGATGATTT 666
Db 1003 CTTGGAAGCCCGAGCCCGAGACCCGGAAGAGCCGCTTGGCTTCACTGATGATTT 1062
Oy 667 GTGGAGACACTGCTGTGGAGATGACAAAGATGCCCGCATTCACGCTGCGGGGCTAAG 726
Db 1063 GTGGAGACACTGCTGTGGAGATGACAAAGATGCCCGCATTCACGCTGCGGGGCTAAG 1122
Oy 727 CCTTACCTCTTAACAGTATGAGCAGACGACCAAGGCTTCAAGCAGCCCAAGCATCCGC 786
Db 1123 CCTTACCTCTTAACAGTATGAGCAGACGACCAAGGCTTCAAGCAGCCCAAGCATCCGC 1182
Oy 787 AATCTGCTGAGCTGGTGGTGTGACTCGGCTAGTATGATCCTGGGGTCAAGGAGGAGGCC 846
Db 1183 AATCTGCTGAGCTGGTGGTGTGACTCGGCTAGTATGATCCTGGGGTCAAGGAGGAGGCC 1242
Oy 847 CAAATGGGGCCAGTGTGCTGCCAGACCTGCGCAGCTTCCTGCTGGCAGCGGGGCTTC 906
Db 1243 CAAATGGGGCCAGTGTGCTGCCAGACCTGCGCAGCTTCCTGCTGGCAGCGGGGCTTC 1302
Oy 907 AACACCCCTGAGAGCTGGGAGCCCTGACCACTTGAACAGCCCATTTCTTTACCCCTGAG 966
Db 1303 AACACCCCTGAGAGCTGGGAGCCCTGACCACTTGAACAGCCCATTTCTTTACCCCTGAG 1362
Oy 967 GACCTGTGTGAGTGTCCACTTGGCAGACGCTGGGTATGGTGTGATGGGCAACCGTCTGT 1026
Db 1363 GACCTGTGTGAGTGTCCACTTGGCAGACGCTGGGTATGGTGTGATGGGCAACCGTCTGT 1422
Oy 1027 GACCCGGCTCGGAGCTGTGCCATTTGTGAGAGATGATGGGCTCCAGTCAAGCTTCACTGCT 1086
Db 1423 GACCCGGCTCGGAGCTGTGCCATTTGTGAGAGATGATGGGCTCCAGTCAAGCTTCACTGCT 1482
Oy 1087 GGTCAATGAACCTGGTCAATGCTTCAACATGCTCCATGACATCCCAAGCCATGACATCACT 1146
Db 1483 GGTCAATGAACCTGGTCAATGCTTCAACATGCTCCATGACATCCCAAGCCATGACATCACT 1542
Oy 1147 TTGAATGGGCTTTTGAAGACCTCTGCGCATGTCTATGGCCCTGTGTATGGCTCATGTGAT 1206
Db 1543 TTGAATGGGCTTTTGAAGACCTCTGCGCATGTCTATGGCCCTGTGTATGGCTCATGTGAT 1602
Oy 1207 CCTGAGAGACCTGTGCTCCCTGCGCATGTGCGCCCTTCTATGACTGACTTCTTGGCAATGGC 1266
Db 1603 CCTGAGAGACCTGTGCTCCCTGCGCATGTGCGCCCTTCTATGACTGACTTCTTGGCAATGGC 1662
Oy 1267 TATGGGCACTGTCTCTTACAAACAGAGGCTCCATGCTGCTGCTGATGGGAGC 1326
Db 1663 TATGGGCACTGTCTCTTACAAACAGAGGCTCCATGCTGCTGCTGATGGGAGC 1722
Oy 1327 TACAAGAGCAGATGACAAAGGCTAGTCA 1358
Db 1723 GGCAGAGACTATGATGCTGACCGCCAGTCCCA 1754

```

RESULT 2  
 US-09-484-970B-58  
 : Sequence 58, Application US/09484970B  
 : Patent No. 6426186  
 : GENERAL INFORMATION:

```

: APPLICANT: Jones, Karen A.
: APPLICANT: Volkmuth, Wayne
: APPLICANT: Walker, Michael G.
: TITLE OF INVENTION: BONE REMODELING GENES
: FILE REFERENCE: PB-0014 US
: CURRENT APPLICATION NUMBER: US/09/484,970B
: CURRENT FILING DATE: 2000-01-18
: NUMBER OF SEQ ID NOS: 172
: SOFTWARE: PERL Program
: SEQ ID NO: 58
: LENGTH: 3706
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc.feature
: OTHER INFORMATION: Incyte ID No. 6426186 007074.1
: LOCATION: 3634, 3638-3639, 3642-3643, 3647-3648, 3652, 3654-3658, 3664, 3674, 368
: OTHER INFORMATION: a, t, c, g, or other
: US-09-484-970B-58

```

Query Match 20.6% Score 280.2; DB 4; Length 3706;  
 Best Local Similarity 62.3%; Pred No. 1,4e-60;  
 Matches 457; Conservative 0; Mismatches 273; Indels 3; Gaps 1;

```

Oy 626 GACCCCGAAGACCAAGCCCTTGGCTTCACTGATGATTTGTGAGACACTGTGTGG 685
Db 169 GAAGCATTAAGAAAGAACGATTTGTGTCCAGTACCCGATGTGGAAACCATGCTTGTGG 228
Oy 686 CAGATGACAAAGATGGCCGCTTCCAGCTGCGGGCTTAAAGCCCTACCTGCTAACAGTCA 745
Db 229 CAGACAGTGTGATGGCAGAAATTCACGGGAGTGTCTAAAGCATTAATCTTCACGTTGT 288
Oy 746 TGGCAGCAGCAGCAGCAAGGCTTCAAGCAGCAAGCAATCCGATCCTGACCTGTGTGG 805
Db 289 TTTGGTGTGACAGCATTTGTACAAACACCCAGCATTTGTATTTACGTTACCTGTGTGG 348
Oy 806 TGACTGTGCTAGTATCTGTGGGTGAGGGGAGAGGGGCCCAAGTGGGGCCAGTGTG 865
Db 349 TGGTAAAGATCTTGGTCATCCAGATGACAGAGAGGGGCCGGAAGTCCCATGCTG 408
Oy 866 CCCAGACCTGGGAGCTTCTGTGCTGGCAGCGGGGCTTCAACACCCCTGAGAGTGG 925
Db 409 CCCGTACTGTGGGAACTTTTGCACACTGGCAGAGGAGCAACCCCACTGACCGGG 468
Oy 926 ACCGTGACCACTTGAACAGCAGCATTTCTTACCCTGAGACCTGTGTGAGATGCCA 985
Db 469 ATGCAGAGCAGCTATGACAGCAATTTCTTCAACAGAGAGACTTGTGTGGTCCAGA 528
Oy 986 CTTGGCAGACGCTGGGTATGCTGATGTGGCAACCTGTGTGACCCGCTCGAGCTGTG 1045
Db 529 CATGTGATCTCTTGGAGAGCTGCTGATGTGGAAGTGTGTATCCGACAGCAAGCTGCT 588
Oy 1046 CCATTTGGAGATATATGGGCTCCAGCTTCACTGCTCTCTCATGAGAGCTGATG 1105
Db 589 CCGTATATGAAGATATGTTTACAGCTGCCCTTCAACAGCCCAATGATTAAGGCTCAG 648
Oy 1106 TCTTAAACATGTCTGATGACAACTCCAGCATGATGATGATGATGGGCTTTGAGCA 1165
Db 649 TCTTAAACATGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 705
Oy 1166 CCTTGGCAGATGATGAGCCCTGTGATGCTGATGCTGATGATGATGATGATGATGATG 1225
Db 706 AGGATTTCCACATGATGATGGGCTCAATGCTTTCAACCTGAGACAGCCAGCTGTGCTC 765
Oy 1226 CCTGAGTGGCCGCTTCACTGACTTCTGTGAGCAATGGCTATGGGCACTGTCTTGG 1285
Db 766 CTTGAGATGCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 825
Oy 1286 ACAAGCAGAGGCTCATGATGATGCTGCTGATGAGGAGTACAAGAGCAGATGATGAC 1345
Db 826 ACAAGCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 885

```

OY 1346 AGGGGTAGTCGA 1358  
 Db 886 ACCGCACTGCCA 898

## RESULT 3

US-09-130-491-1  
 : Sequence 1, Application US/09130491  
 : Patent No. 6416974  
 : GENERAL INFORMATION:

: APPLICANT: Holtzman, Douglas A.  
 : APPLICANT: Goodheart, Andrew D.J.  
 : TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83

: FILE REFERENCE: 09404/041001

: CURRENT APPLICATION NUMBER: US/09/130,491

: EARLIER FILING DATE: 1998-08-07

: EARLIER APPLICATION NUMBER: US 60/058,108

: EARLIER FILING DATE: 1997-09-05

: EARLIER APPLICATION NUMBER: US 60/054,961

: NUMBER OF SEQ ID NOS: 16

: SOFTWARE: FastSeq for Windows Version 3.0

: SEQ ID NO 1

: LENGTH: 4676

: TYPE: DNA

: ORGANISM: Homo sapiens

: FEATURE:

: NAME/KEY: CDS

: LOCATION: (460)...(3360)

US-09-130-491-1

Query Match 20.6%; Score 280.2; DB 4; Length 4676;  
 Best Local Similarity 62.3%; Pred. No. 1.6e-60;  
 Matches 457; Conservative 0; Mismatches 273; Indels 3; Gaps 1;

OY 626 GACCCCGAAGACCAAGCGCTTGGCTTCTGCTAGTAGATTGTGGAGACACTGTGTGG 685  
 Db 1196 GAAGCATTAAGAAAGCAAGATTGTGTCTCAGTCAACCGCTATGTGMAACCATGCTTGG 1255  
 OY 686 CAGATGACAAAGTGGCCGCTTCCAGCGTGGGGCTAAAGCGTACGTAACAGGA 745  
 Db 1256 CAGACCACTCATGGCAGAAATTCAGCGCAGTGTCTAAAGCATTAACCTTCCAGCTTGT 1315  
 OY 746 TGGCAGCAGCAAGCCAAAGCCTTCAAGCACCACCAAGCTCCGATCCTGTGAGTGG 805  
 Db 1316 TTTGGGTGGCAGCGAGATTGTACAAACACCCAGCATTCGTAATTCATTAGCCGGTGG 1375  
 OY 806 TGACTCGGCTAGTATCTCTGGGGTCAGCGAGAGAGGGCCCCCAAGTGGGGCCAGTCTG 865  
 Db 1376 TGGTGAAAGATCTTGGTCAATCCAGATGAACAGAAAGGGCCGGAAGTGAATCCAAATGCTG 1435  
 OY 866 CCCAGACCTGGGCGCTTGTGTGCTGGCAGCGGGGGCTCAACGCCCTGAGCACTGG 925  
 Db 1436 CCGTCACCTCTGGGAACTTTTGCAACTGGCAGAACGACACCAACCCCACTGACCGGG 1495  
 OY 926 ACCCTGACCACTTGGACAGAGCATTCGTTTACCCTGAGGACCTGTGTGAGTCTCCA 985  
 Db 1496 ATGCAGACACTATGACACAGAAATTTCTTTCACAGACAGACACTGTGTGGTCCAGA 1555  
 OY 986 CTTGGCAGACGCTGGGTATGGCTGATGGGCAACCGTCTGTGACCCGGCTGGAGCTGTG 1045  
 Db 1556 CATGTGATACCTTGGGATGGGTGATGGGAATGTGTGATCCGAGCAAGAACCTGCT 1615  
 OY 1046 CCAATGTGGAGGATGGGGCTCGAGCTTCACTGCTCCTCATCAACTGAGCTGCATG 1105  
 Db 1616 CCGTATATGAAAGATATGATTTTACAAGCTGCCCTTACACACACCCCATGATTAAGCCAG 1675  
 OY 1106 TCTTCAACATGCTCCATGACAACTCCAAAGCCATGCTGATGATGATGGGCTTTGAGA 1165  
 Db 1676 TGTTAAACATGACACATATGATGCAAAAGACAGTGTGCCAGCTTAATGG - - -TGTGAACC 1732  
 OY 1166 CCTTGGCCATGTCAATGGCCCTGTGATGGCTCATGTGGATCTCTAGGAGCCCTGTGCTCC 1225

Db 1733 AGGATTCACCATGATGAGCGCTCAATGCTTTCCAACTGGACACAGCCAGCTTGGTCTC 1792  
 OY 1226 CCTGACAGTCCCGCTTCACTGACTCTCTGACAAATGGATGGCACTGCTCTAG 1285  
 Db 1793 CTTGCAAGTCCCTACATGATGATCATCATTTCTGGATTAATGTGATGGGAAATGTTGATGG 1852  
 OY 1286 ACAAAACAGAGGCTTCCATTTGATGCTGCTGTGAGTGGGAGTACAGAGCAGATGACA 1345  
 Db 1853 ACAAGCTTCAGATATCCCATPACAGGCTCCCAAGCGATCTCCCTGGACACCTGTACATGCCA 1912  
 OY 1346 AGGGGTAGTCGA 1358  
 Db 1913 ACCGCACTGCCA 1925

## RESULT 4

US-09-392-184-1  
 : Sequence 1, Application US/09392184  
 : Patent No. 6395889  
 : GENERAL INFORMATION:

: APPLICANT: Robison, Keith E.  
 : TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN

: FILE REFERENCE: 5800-55

: CURRENT APPLICATION NUMBER: US/09/392,184

: EARLIER FILING DATE: 1999-09-09

: NUMBER OF SEQ ID NOS: 33

: SOFTWARE: FastSeq for Windows Version 3.0

: SEQ ID NO 1

: LENGTH: 4858

: TYPE: DNA

: ORGANISM: Homo sapiens

: FEATURE:

: NAME/KEY: misc.feature

: OTHER INFORMATION: repolysin (ADAM family of metalloprotease)

: LOCATION: (1)...(4858)

: NAME/KEY: misc.feature

: LOCATION: (1)...(4858)

: OTHER INFORMATION: n = A,T,C or G

US-09-392-184-1

Query Match 19.9%; Score 270.8; DB 4; Length 4858;  
 Best Local Similarity 62.4%; Pred. No. 3.5e-58;  
 Matches 458; Conservative 0; Mismatches 272; Indels 4; Gaps 2;

OY 626 GACCCCGAAGACCAAGCGCTTGGCTTCTGCTAGTAGATTGTGGAGACACTGTGTGG 685  
 Db 1190 GAAGCATTAAGAAAGCAAGATTGTGTCTCAGTCAACCGCTATGTGMAACCATGCTTGG 1249  
 OY 686 CAGATGACAAAGTGGCCGCTTCCAGCGTGGGGCTAAAGCGTACCTGCT -AACAGTG 744  
 Db 1250 CAGACCACTCATGGCAGAAATTCAGCGCAGTGTCTAAAGCATTAACCTTCCAACTTG 1309  
 OY 745 ATGCAGCAGCAAGCCAAAGCCTTCAAGCACCACCAAGCATCCGATCTGTGAGTGGTG 804  
 Db 1310 TTTTGGTGGGACGCCAGATTGTGACAAACCCCACTTCGTAATTCAGTTAGCTGGTG 1369  
 OY 805 GTGACTCGGCTAGTATCTCTGGGTCAGCGAGGGGGGCCCAAGTGGGGCCCACTGCT 864  
 Db 1370 GTGTGAAGATCTTGGTCAATCCAGATGAACAGAAAGGGCCGGAAGTGAATCACTCAATCT 1429  
 OY 865 GCCCAGACCTTCGCGAGCTTGTGTGCTGGCAGCGGGGCTCAACACCCCTGAGGACTCG 924  
 Db 1430 GCCCTCATCTCTGGGAACTTTTGCAACTGGCAGGAAGACAGCAACACCCCACTGACCGG 1489  
 OY 925 GACCTGACCACTTTGACACAGCCATTCGTTTACCCTGAGGACCTGTGTGAGTCTCC 984  
 Db 1490 GATGCAGAGCACTATGACACAGCAATTTCTTTCACAGACAGGACTGTGTGGGTCCAG 1549  
 OY 985 ACTTGGACACGCTGGGTATGGCTGATGGGCAACCGTCTGTGACCCGGCTGGAGCTGT 1044  
 Db 1550 ACATGTGATACCTTGGGATGGGTGATGTTGGAAGTGTGTGATGCAAGAGAGCTGC 1609

QY	634	AGACCCAGCGCTTCCTGCTACGAGTAAGTTGTGTGACACACTGATGTGGCAATATAC	693
	111	111	111
Pb	950	AGAGCAAGAGGTTGTGTGCCGAGGCTCGCTTGTTGGAAACACTTCTGTGGTCTGATCGG	1009
QY	694	AAGATGCCCCCATTTCCACGGTGGCGGGGCTTAAAGCGTTACTCTTAACATGATGGCAGCA	753
	111	111	111
Pb	1010	TCCATGCTGCTCTTCTATTGTGGACCGACTCGACACACACATCTCTACGGTATGTCAATG	1069
QY	754	GCACCCAGGAGCGCTTCAAGCACGCCCAAGCATCCGCAATCTGTGAGCTTGATGACTGG	813
	111	111	111
Pb	1070	GCACCCCGGAATCTACAAAGCACCCGACACTAGGAATCTCCGTACACCTTGATGGTGTGAAA	1129
QY	814	CTAGTGAATCCTGGGGGTAGCGAGAGAGGGGCCCAAGTGGGCGCCAGTCTGCCAGACC	873
	111	111	111
Pb	1130	GTCGTAATAGTGGAAMAAAGAAATAGTGGGGCCGGGAATCTCCGACAAAGGGGGCGCTCA	1189
QY	874	CTGGCGAGCTTCTGTGCTGTGGCAGCGGGGCGCTCAACACCCCTGAAGACTCGGACCTTAC	933
	111	111	111

Oy	632	GAAGGCCAACGGCTTTGGCTTCACTAGTGTGATTTTGGAGACAGTGGTGGGAGATG	691
Db	2653	GTAGGACCAACGGCTTGTGTCTGTGAGGCGCGCTCTTGAGAACGCTGCTGGTGGCCATG	2594
Oy	652	ACAAGATGGCGCCATTCACAGGTGCGGGGCTAAAGCGTACTGCTAACATGATGACAG	751
Db	2593	CGTCCATAGTGGCTCTTACGGGGCGGACCTGCAGAACCACTCTGACGTTAAATGTCG	2534
Oy	752	CAGCAGCCAAAGCGCTTCAACGACCCCAAGCATCCGCAATCTGTACAGTGGTGGTGAATC	811
Db	2533	TGGCAGCCCGGATTTACAGACGCCCAAGCATCAAGAAATTCATCACTGATGGTGGTAA	2474
Oy	812	GCGTAGTGATCTTGGGGTTCAGGCGAGAGAGGGGCCCAAGTGGGGCCCACTGCTGCCAGA	871
Db	2473	AAGTGTGATCGTGAAGATGAAGAAATATGGGGCCCAAGGCTGTCCGACATGGGGGGCTTA	2414
Oy	872	CGCTGGCAGATTTGTGGCTTGGGACAGGGGGCTTCAAACCCCTGAGGCTGGAGCCCTG	931
Db	2413	CACATCGTAACCTTTGTGCACTGGCAGCGGGCTTTCAACGACGCCACGACGCCGCCACAG	2354





```

Db 1122 GATTATGAGGAGCATTCATCATGTGACACCCCTGGGAATGGAGAGTTGGGACCATATGT 1181
Qy 1027 GACCCGCTCGGAGCTGCGCATTTGTGGAGATGATGGGCTCCAGTCCAGCTTCACTGCT 1086
Db 1182 TCTCCGAGACGAGCTGTGCATGATGAAGATGATGGCTTCATGACACCTTCACTGCTG 1241
Qy 1087 GCTCATGAATGGGTGATGCTTCAATGCTTCATGACACCTCAAGCATGATCATGAT 1146
Db 1242 GCTCATGAATGGGTGATGCTTCAATGCTTCATGACACCTCAAGCATGATCATGAT 1298
Qy 1147 TTGATGGGCTTGAAGACCTCTGCGCATGATGAGGCTTGTATGGCTCATGATGATGAT 1206
Db 1299 GAGAACTTGGGATGACAGACAGCAAGGTTAATGCTTCAATGCTTCAAGCATGAT 1358
Qy 1207 CCTGAGAGACCTTGGTCCCTCCAGTGGCCGCTTCATGATGATGCTTCAAGCATGAT 1266
Db 1359 GCATCCAGACCTTGGTCCCAATGACAGCTGACGACCATGACAGATTCCTGATGATGAT 1418
Qy 1267 TATGGGCACTGCTCTTGAACAAACCA 1293
Db 1419 CATGCTAATGTTGCTGACCTACCA 1445

```

## RESULT 9

```

US-09-122-126B-14
: Sequence 14, Application US/09122126B
: Patent No. 6451575
: GENERAL INFORMATION:
: APPLICANT: Bristol-Myers Squibb Company
: TITLE OF INVENTION: ASGRECAN DEGRADING METALLO PROTEASES
: FILE REFERENCE: DM6909
: CURRENT APPLICATION NUMBER: US/09/122.126B
: CURRENT FILING DATE: 1998-07-24
: NUMBER OF SEQ ID NOS: 21
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 14
: LENGTH: 3250
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (121)..(2910)
US-09-122-126B-14

```

```

Query Match 12.9%; Score 175.8; DB 4; Length 3250;
Best Local Similarity 55.5%; Pred. No. 1.3e-34;
Matches 361; Conservative 0; Mismatches 287; Indels 3; Gaps 1;

```

```

Qy 643 CGCTTTGCTTCACTAGTAGATTTGTGGAGACACTGTGTGGATGACAAATGGCC 702
Db 901 CGCTTCATCTCCCGGCGCCGCGAGGTGAGCTTGTGTGGCTGACGCGCTCATGGCG 960
Qy 703 GCATTCACGGGGGGGCTAAAGCGCTAAGCTGATGATGAGAGACAGACCCAG 762
Db 961 CGGTGTATGCGCGGGGCTGACGATTAACCTGCTGACCTGCGCTCATGGCAATAG 1020
Qy 763 GCGTTCAAGCAACCAAGCATCGCAATCTGTCACTGTGTGTGTGACTCGCTACTGATC 822
Db 1021 CTGTACAGCAATGTAGATGAGAACCAATCCGCTGGCGGTGTGAAGTGTGTG 1080
Qy 823 CTGGGGTCAAGGAGAGAGGGGCCCCAATGGGGCCCACTGCTGCCAAGACCTGGCGAGC 882
Db 1081 CTAGGCGCAAGAGCAAGAGCGCTGAGAGTGAAGCAAGACGCTGCCACCACTCAAGAAC 1140
Qy 883 TTCTGTGCTGACGAGGCGGCTCAACACCCCTGAGAGCTGGAGCCCTGACCACTTTGAC 942
Db 1141 TTTTCAAGTGGACACACACACACACAGCTGGAGATGACACATGAGAGACACTAGCAT 1200
Qy 943 ACAGCATCTTGTATACCGCTGACAGACCTGTGTGAGTCTCCACTTGGAGACAGCTGGGT 1002
Db 1201 GCAGTATCTGTTTACGCGGAGAGATTTATGTGGCATCATTCATGTGACACACCTGGGA 1260

```

```

Qy 1003 ATGCTGATGTGGGACCGCTGTGACCCGCTCGAGCTGTGCCATTGTGGAGATGAT 1062
Db 1261 ATGCGAGAGCTTGGAGACCATATGTTCTCCAGAGCGCACCTGTGCTGATTTGAAGACGAT 1320
Qy 1063 GGGCTCCAGTACGCTTACGCTGCTCATGACAGTGGGTCAATGCTTGAACATGCTCAT 1122
Db 1321 GGGCTCCAGGAGCGCTTACGCTGCTGCTGACGAAATGAGCAATTAATGCTGCTCCAT 1380
Qy 1123 GACAACCTCAAGCCATGCAATGATGATGAGGCGCTTGGACACCTCGCGCATGATCATG 1182
Db 1381 GACGATTCGAA---ATTCTGGAAGAGACCTTTGGTTCCACAGAAATGAAGCGTTAATG 1437
Qy 1183 GCGCTGTGATGCTCATGATGATGATCTGAGAGAGCCCTGCTCCGAGTGGCCGCTTC 1242
Db 1438 TCTTCATCTCTTACGAGCATGATGATCATCTAAGCCCTGCTGCAAAATGACATTCAGCCAC 1497
Qy 1243 ATCACTGACTTCTGAGACAAATGATGGGACATGCTCTTGAACAAACCA 1293
Db 1498 ATCAGCAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1548

```

## RESULT 10

```

US-09-392-184-5/C
: Sequence 5, Application US/09392184
: Patent No. 6395889
: GENERAL INFORMATION:
: APPLICANT: Robison, Kelch E.
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN
: FILE REFERENCE: 5800-55
: CURRENT APPLICATION NUMBER: US/09/392.184
: CURRENT FILING DATE: 1999-09-09
: NUMBER OF SEQ ID NOS: 33
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 5
: LENGTH: 5357
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: (1)..(5357)
: OTHER INFORMATION: repolysin (ADAM family of metalloprotease)
: NAME/KEY: misc.feature
: LOCATION: (1)..(5357)
: OTHER INFORMATION: n = A,T,C or G
US-09-392-184-5

```

```

Query Match 10.1%; Score 137.4; DB 4; Length 5357;
Best Local Similarity 54.6%; Pred. No. 5.7e-25;
Matches 348; Conservative 0; Mismatches 271; Indels 18; Gaps 3;

```

```

Qy 667 GTGGAGACACTGGGTGGAGATGACAAATGGCCGCTTCCAGCG--TGCGGGGCTA 723
Db 4589 GTGGAGACCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4530
Qy 724 AAGCGTACTGCTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 783
Db 4529 GAGAGCTATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4470
Qy 784 GCGAATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 843
Db 4469 GCGAATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4410
Qy 844 CCGCAAGTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 903
Db 4409 CTAAGATGACAGCAACATGACAGACACCTGGAAGCTTCTGCAAGTGGGCAAAAAGC 4350
Qy 904 CTCAACACCCCTGAGAGACTGAGACCTGACACCTTTGACACAGCAATTTGTTAACCCCT 963
Db 4349 ATCAAGATGAGAGGGGATGCCATTCCTGACCATGACATGCTGCTGCTGCTGCTGCTGCT 4290
Qy 964 CAGAACCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1017

```

```

Db 4289 AAGGACCTCTGTGACAGCATGAAACCGGCCCTGTGAGACCCCTGGACCTGCCATGTGGGC 4230
Oy 1018 ACCGCTGTGACCGCGCTGGAGCTGTGCATTTGAGAGATGATGGCTCCAGTCAGCC 1077
Db 4229 GGCATGTGACCGCGCTGGAGCTGTGCATTTGAGAGATGATGGCTCCAGTCAGCC 4170
Oy 1078 TTCACCTGTGCTGATGAACTGGGTCATGTCTTCAACATGCTCCATGACAACTCAAGCCA 1137
Db 4169 TTCACCTGTGCTGATGAACTGGGTCATGTCTTCAACATGCTCCATGACAACTCAAGCCA 4119
Oy 1138 TGCATCATGTTGATGAGGCGCTTTGAGCACCTCTCCCATGTATGAGCCCTGTGATGCT 1197
Db 4118 ACGGCGCATGCTGTGAGAGCGCTGTCCCTGAGTGCCTTCATGCTGACATGCTCTG 4059
Oy 1198 CATGCGATCTGTGAGAGCGCTGTCCCTGAGTGCCTTCATGCTGACATGCTCTG 1257
Db 4058 TACGACGCGCGCTCCCTCAGCTGTGCTCCGCTGACGCGCGCATATATACACAGTTCTT 3999
Oy 1238 GACAACTGCTATGAGGCGCTGTCTCTTGAACAACGAG 1294
Db 3998 GACCGTGGGTGGGCGCTGTGCTGAGACGACCTCTG 3962

```

```

RESULT 11
US-09-369-364A-6
: Sequence 6, Application US/09369364A
: Patent No. 6391610
: GENERAL INFORMATION:
: APPLICANT: Apte, Suneel
: APPLICANT: Hurskainen, Tiina L.
: APPLICANT: Hirohata, Satoshi
: TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
: FILE REFERENCE: 26473/4007/10-30-00
: CURRENT APPLICATION NUMBER: US/09/369,364A
: NUMBER OF SEQ ID NOS: 31
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 6
: LENGTH: 3218
: TYPE: DNA
: ORGANISM: Homo sapiens ADAMTS-7
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (13)..(3003)
US-09-369-364A-6

```

```

Query Match          9.9%; Score 134.4; DB 4; Length 3218;
Best Local Similarity 54.4%; Pred. No. 2,7e-24;
Matches 345; Conservative 0; Mismatches 271; Indels 18; Gaps 3;

Oy 670 GAGACACTGTGTGAGCATGACAGATGAGCGCATTCACAG--TCCGGGGCTAAAG 726
Db 745 GAGACCTGTGTGAGCATGACAGATGAGCGCATTCACAG--TCCGGGGCTAAAG 804
Oy 727 GCGTACTGTCTAACAAGTGTGAGCAGCAGCAGCGCAAGCGCTTCAAGCAGCCAGCATCCGC 786
Db 805 ACCTATGTCTGACCATCATGTAACATGAGTGTGCTGCGCTTTTCATGACCCAGCATTTGAG 864
Oy 787 AATCTGTAGCTGTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 846
Db 865 AACCCCATCATCATCAACCATGTGTGCGCTGTGCTGTGTAAGATGAGAGGAGACCTA 924
Oy 847 CAAGTGGGGCCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 906
Db 925 AAGATCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 984
Oy 907 AACACCCCTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 966
Db 985 AACATGAAGGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1044
Oy 967 GACCTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1020
Db 1045 GACCTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1104

```

```

Oy 1021 GTCGTGACCGCGCTGGAGCTGTGCATTTGAGAGATGATGGCTCCAGTCAGCCCTTC 1080
Db 1105 ATGTGCGACCGCGCTGGAGCTGTGCATTTGAGAGATGATGGCTCCAGTCAGCCCTTC 1164
Oy 1081 ACTGCTGTGATGAACTGGGTCATGTCTTCAACATGCTCCATGACAACTCAAGCATGC 1140
Db 1165 ACTGTAGCCCGACGACCTGGGCGCATGTTTGGCATTCAGCATGACGG-----AAGC 1215
Oy 1141 ATCATTTGATGAGGCGCTTTGAGCACCTCTCGCCATGTATGAGCCCTGTATGCTCAT 1200
Db 1216 GGCATGACTGTGAGCGCGCTGTGGAAACGACCTTTCATGATGCTGACAGCTCTGTAG 1275
Oy 1201 GTGATCTGTGAGAGCGCGCTGTGGAAACGACCTTTCATGATGCTGACAGCTCTGTAG 1260
Db 1276 GACGCGCGCTCCCTCAGCTGTGCTCCGCTGACAGCGCGCATATATACAGGTTCTTAC 1335
Oy 1261 AATGCTATGAGGCGCTGTCTTGAACAACGAG 1294
Db 1336 CGTGGTGGGCGCTGTGCTGAGACGACCTCTG 1369

```

```

RESULT 12
US-09-392-184-6
: Sequence 6, Application US/09392184
: Patent No. 6395889
: GENERAL INFORMATION:
: APPLICANT: Robison, Keith E.
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN
: FILE REFERENCE: 5800-55
: CURRENT APPLICATION NUMBER: US/09/392,184
: NUMBER OF SEQ ID NOS: 33
: SOFTWARE: FASTSEQ for Windows Version 3.0
: SEQ ID NO 6
: LENGTH: 703
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: (1)...(703)
: OTHER INFORMATION: reprotolysin (ADAM family of metalloprotease)
: NAME/KEY: misc.feature
: LOCATION: (1)...(703)
: OTHER INFORMATION: n = A,T,C or G
US-09-392-184-6

```

```

Query Match          9.7%; Score 132.4; DB 4; Length 703;
Best Local Similarity 65.0%; Pred. No. 5,1e-24;
Matches 212; Conservative 0; Mismatches 111; Indels 3; Gaps 1;

Oy 988 TCGCAGACCTGTGTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 1047
Db 14 TCGCAGACCTGTGTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 73
Oy 1048 ATTGGAGAGATGAGGCTTCCAGTCACTTCATGCTGTGCTATGATGATGATGATGATGAT 1107
Db 74 GTGATGAGAGATGAGGCTTCCAGTCACTTCATGCTGTGCTATGATGATGATGATGATGAT 1133
Oy 1108 TTCACATGCTCCATGACAACTCAAGCCATCATGATTTGAAGGGGCTTTGAGCAGC 1167
Db 134 CTCAGATGCTCCAGAGACAGCTCAAGCCCTTCACACGGCTCTTGGGGCCATGGGGA-- 191
Oy 1168 TCTGCCATGATATGAGGCGCTGTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1227
Db 192 -AGCACACATGATGACACCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 250
Oy 1228 TCGAGTGGCGCTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1287
Db 251 TCGAGGCGCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 310
Oy 1288 AACGAGGCTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1313

```



```
Db 1158 GTGTTAACATGCTCCTCATGATGACACAA-----CAATGTAAAGAAAGAGAGTTAAG 1211
OY 1165 ACCCTCCGATGATGAGCCCTGTGATGCTCATGTGATCTGAGAGAGCCCTGGTCC 1224
Db 1212 ACTGCCAGACATGATGATGCTGCTCAACACTGAACTTCTACACCAACCCCTGGATGGTCA 1271
OY 1225 CCCTGCAGTCCCGCTTGTATCATGACTGACTGACAAATGCTATGGCGACTGTCTTAA 1284
Db 1272 AAGTGTAGTCGAAATATATATCACTAGAGTTTAAACACACTGGTTATGGCGAGTGTTCCTT 1331
OY 1285 GACAACCCAGA 1295
Db 1332 AACGAACCTGA 1342
```

## RESULT 15

```
US-09-392-184-23/C
; Sequence 23, Application US/09392184
; Patent No. 6395889
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: PROTEASE HOMOLOGS
; FILE REFERENCE: 3800-55
; CURRENT APPLICATION NUMBER: US/09/392,184
; CURRENT FILING DATE: 1999-09-09
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 385
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(385)
; OTHER INFORMATION: pcnp (procollagen N-proteinase)
US-09-392-184-23
```

Query Match 4.6%; Score 62.8; DB 4; Length 385;

Best Local Similarity 65.9%; Pred. No. 8.7e-07;

Matches 91; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

```
OY 1004 TGGCTGATGTGGGACCGCTGTGTGACCCGCTGGAGCTGTGCCATTTGTGAGAGATGATG 1063
Db 385 TGGCTGAACTGGGAGCAATTTGTATCCCTATAGAAAGCTGTCTATAGTGAAGATAGTG 326
OY 1064 GGCTCCAGTCACGCTTCTACTGCTGCTCATGAACTGGGTCATGTCTTCAACATGCTCATG 1123
Db 325 GATTGAGTACAGCTTTTACGATGCCCATGAGCTGGGCCATGTGTTAAATGCTCATG 266
OY 1124 ACACTCCAAAGCATGCA 1141
Db 265 ATGACACACAAATGTA 248
```

Search completed: April 26, 2003, 05:36:01  
Job time : 113.23 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 25, 2003, 07:28:33 ; Search time 596.069 Seconds  
(without alignments)  
5134.413 Million cell updates/sec

Title: US-10-050-200-1

Perfect score: 1359

Sequence: 1 gaattcgcatgtccagac.....atgacaagggtagtcgac 1359

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database : N\_Geneseq\_101002.\*

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1320	97.1	4301	22	AA168146
2	1320	97.1	4303	21	AA168146
3	1318.8	97.0	4192	20	AA168146
4	1316.8	96.9	4406	22	AA168146
5	1316.8	96.9	4407	21	AA168146
6	1316.8	96.9	4407	22	AA168146
7	543.4	40.0	1878	24	AA168146
8	281	20.7	4712	20	AA168146
9	280.2	20.6	2853	22	AA168146

10	280.2	20.6	3261	20	AA168146	Human MET11 encodi
11	280.2	20.6	3261	19	AA168146	Human MET11 coding
12	280.2	20.6	4014	22	AA168146	Human integrin lig
13	280.2	20.6	4676	20	AA168146	Human Tango-71 enc
14	279.2	20.5	2853	22	AA168146	Human metalloprote
15	279.2	20.5	2853	24	AA168146	Human metalloprote
16	279.2	20.5	2930	24	AA168146	Human protease cdn
17	279.2	20.5	2937	24	AA168146	Human protease prt
18	279.2	20.5	3446	24	AA168146	Human metalloprote
19	277.4	20.4	2184	20	AA168146	Human protease cdn
20	276.4	20.3	4180	20	AA168146	Human ADAMTS-1 cod
21	276.4	20.3	4180	22	AA168146	Human MET11 relate
22	276	20.3	1518	22	AA168146	D67076 cDNA clone.
23	270.8	19.9	4858	24	AA168146	Rat metalloprote
24	251	18.5	2346	24	AA168146	Human protease cdn
25	249.8	18.4	3008	20	AA168146	Human metalloprote
26	249.8	18.4	3008	22	AA168146	Human MET12 encodi
27	248.2	18.3	2670	22	AA168146	Human MET12 coding
28	248.2	18.3	2670	22	AA168146	Human ADAM-type me
29	246.4	18.1	3638	22	AA168146	Human ADAMTS-8 CD
30	233.8	17.2	3126	24	AA168146	Human ADAMTS-8 CDN
31	228.4	16.8	739	22	AA168146	Human ADAMTS-8 CDN
32	227.8	16.8	1143	21	AA168146	Human metalloprote
33	219.2	16.1	1104	24	AA168146	Human metalloprote
34	180.6	13.3	3002	22	AA168146	Human protease cdn
35	178.8	13.2	652	22	AA168146	Murine ADAMTS-5 CD
36	175.8	12.9	3250	20	AA168146	Human ADAMTS-5 CD
37	175.8	12.9	5530	21	AA168146	Human aggrecan deg
38	152.2	11.2	2751	21	AA168146	Human ORF990
39	147	10.8	528	22	AA168146	Rat metalloprote
40	143.6	10.6	9120	22	AA168146	AA286889 cDNA clon
41	143.6	10.6	9120	22	AA168146	Human nervous syst
42	137.8	10.1	9248	22	AA168146	Human nervous syst
43	137.8	10.1	9248	22	AA168146	Human MET11 relate
44	137.4	10.1	5061	22	AA168146	AB001735 cDNA clon
45	137.4	10.1	5353	22	AA168146	Human ADAM-type me

## ALIGNMENTS

RESULT 1	AA168146	standard; cDNA; 4301 BP.
ID	AA168146	
AC	AA168146	
XX		
DT	13-DEC-2001	(first entry)
XX		
DE	Human aggrecanase-1 encoding cDNA seq ID NO 2.	
XX		
KW	Human; aggrecanase-1; ADAMTS4; promoter; antiarthritic; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	401..2914
FT		/*tag- a
FT		/product= "Aggrecanase-1"
XX		
PN	JP2001245663-A.	
PD	11-SEP-2001.	
XX		
PE	06-MAR-2000; 2000JP-0059952.	
XX		
PR	06-MAR-2000; 2000JP-0059952.	
XX		
PA	(RISA ) EISAI CO LTD.	
XX		
DR	WPI: 2001-610072/70.	
XX		
DR	P-PSDB: AAG78228.	
XX		

PT New DNA for use as a promoter for controlling the expression of  
 PT aggreganase-1 -  
 XX

PS Example 1 : Page 5-9; 12pp; Japanese.

CC The invention relates to a DNA acting as a promoter controlling the  
 CC expression of aggreganase-1 (ADAMTS4) resulting in antiarthritic  
 CC activity. The DNA is used in a method for screening a compound affecting  
 CC the expression control of the aggreganase-1 gene.  
 XX

SO Sequence 4301 BP; 844 A; 1267 C; 1232 G; 958 T; 0 other;

Query Match 97.1%; Score 1320; DB 22; Length 4301;  
 Best Local Similarity 98.5%; Pred. No: 0;

Matches 1332; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 7 GCCATGTCCCAAGAGGCTGCATCCCGGAGGGCTTGGCAGGGCGCTGGCTGGGGA 66  
 DB 398 GCCATGTCCCAAGAGGCTGCATCCCGGAGGGCTTGGCAGGGCGCTGGCTGGGGA 457  
 QY 67 GCCCAACCTGCTGCTGCTCCCATGTGCGGCTGCTGCTGCTGCTGCTGCTGCTG 126  
 DB 458 GCCCAACCTGCTGCTGCTCCCATGTGCGGCTGCTGCTGCTGCTGCTGCTGCTG 517  
 QY 127 CTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 186  
 DB 518 CTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 577  
 QY 187 GAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 246  
 DB 578 GAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 637  
 QY 247 AGCATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 306  
 DB 638 AGCATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 697  
 QY 307 TCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 366  
 DB 698 TCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 757  
 QY 367 GGTGGAGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 426  
 DB 758 GGTGGAGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 817  
 QY 427 GCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 486  
 DB 818 GCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 877  
 QY 487 CTCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 546  
 DB 878 CTCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 937  
 QY 547 ATCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 606  
 DB 938 ATCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 997  
 QY 607 CTTGGAAGCCCGAGCCCGAGCCCGAGAGAGCCGCTTTCCTTCACTGAGTAGATT 666  
 DB 998 CTTGGAAGCCCGAGCCCGAGCCCGAGAGAGCCGCTTTCCTTCACTGAGTAGATT 1057  
 QY 667 GTGGAGACACTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 726  
 DB 1058 GTGGAGACACTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1117  
 QY 727 CGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 786  
 DB 1118 CGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1177  
 QY 787 AATCTGTGACCTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 846  
 DB 1178 AATCTGTGACCTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1237  
 QY 847 CAAGTGGGGCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 906

DB 1238 CAAGTGGGGCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1297  
 QY 907 AACACCCCTGAGAGCTCGGACCTTGACACACGCTTGTACCGCTGAG 966  
 DB 1298 AACACCCCTGAGAGCTCGGACCTTGACACACGCTTGTACCGCTGAG 1357  
 QY 967 GACCTGTGTGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1026  
 DB 1358 GACCTGTGTGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1417  
 QY 1027 GACCCGCTCGGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1086  
 DB 1418 GACCCGCTCGGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1477  
 QY 1087 GCTCATGAACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1146  
 DB 1478 GCTCATGAACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1537  
 QY 1147 TTGAATGGGCTTTGAGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1206  
 DB 1538 TTGAATGGGCTTTGAGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1597  
 QY 1207 CCTGAGAGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1266  
 DB 1598 CCTGAGAGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1657  
 QY 1267 TATGGGCACTGTCTG 1326  
 DB 1658 TATGGGCACTGTCTG 1717  
 QY 1327 TACAAGACAGACATGACAAGGCTAGTCA 1358  
 DB 1718 GGCAAGACTATGATGCTGACCCGAGTGCA 1749

## RESULT 2

AAA95826  
 ID AAA95826 standard; cDNA; 4303 BP.

XX AAA95826;

XX 23-FEB-2001 (first entry)

XX Human metalloproteinase KIAA0688 cDNA.

DE Human: KIAA0688; metalloproteinase domain; thrombospondin domain;

KW a disintegrin and metalloproteinase domain; ADAM;

KW vaccine; neurotrophic; neuroprotective; antiparkinsonian;

KW cerebroprotective; cytoskeletal; antiarthritic; immunosuppressive;

KW Alzheimer's disease; Parkinson's disease; stroke; cancer; arthritis;

KW autoimmune disease; brain tumour; brain injury; ss.

OS Homo sapiens.

XX Key

XX Location/Qualifiers

XX CDS 400..2914

XX FT /tag= a

XX FT /product= "KIAA0688"

XX WO200053774-A2.

XX 14-SEP-2000.

XX 08-MAR-2000; 2000WO-0506237.

XX 08-MAR-1999; 99US-0264585.

XX (NEUR-) NEUROCRINE BIOSCIENCES INC.

XX PA Keiner GS, Clark M, Maki RA;

XX PI WPI; 2000-594326/56.

XX DR



DR P-PSDB; AAB21256.  
 XX Polyucleotide encoding novel members of a disintegrin,  
 PT metalloproteinase and thrombospondin domain protein family used to  
 PT prevent and treat Alzheimer's disease, cancer and autoimmune diseases -  
 XX  
 PS Example 3; Fig 11; 129pp: English.  
 XX  
 CC The present sequence encodes human metalloproteinase KIAA0688. KIAA068  
 CC was identified by as a member of the ADAMTS family of proteins by  
 CC searching Genbank for sequences similar to ADAMTS-1 and ADAMTS-3. The  
 CC ADAMTS family of proteins contain thrombospondin, integrin and  
 CC metalloproteinase domains. ADAMTS polypeptides are useful for the  
 CC manufacture of medicaments for treating conditions associated with  
 CC neuroinflammation and/or neurodegeneration, such as Alzheimer's disease,  
 CC Parkinson's disease and stroke. They are also useful for treating  
 CC conditions associated with cell proliferation, cell migration,  
 CC inflammation and/or angiogenesis, such as cancer, arthritis and  
 CC autoimmune diseases. They can be used to treat patients afflicted with an  
 CC invasive tumour, a brain tumour or brain injury.  
 XX  
 SQ Sequence 4303 BP; 844 A; 1267 C; 1233 G; 958 T; 1 other:  
 Query Match 97.1%; Score 1320; DB 21; Length 4303;  
 Best Local Similarity 98.5%; Pred. No. 0;  
 Matches 1332; Conservative 0; Mismatches 20; Indels 0; Gaps 0;  
 OY 7 GCCATGTCAGAGAGAGGCTGATCCGGAGAGGGCTGGAGGGGCTGGCTGGGGGA 66  
 Db 398 GCCATGTCAGAGAGAGGCTGATCCGGAGAGGGCTGGAGGGGCTGGCTGGGGGA 457  
 OY 67 GCCCAACCCCTGCTGCTGCTCCCATTTGGCGGCTCTGCTGGTGGTGGCTGGCTG 126  
 Db 458 GCCCAACCCCTGCTGCTGCTCCCATTTGGCGGCTCTGCTGGTGGTGGCTGGCTG 517  
 OY 127 CTACTGCTGGGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 186  
 Db 518 CTACTGCTGGGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 577  
 OY 187 GAGATGCTGTTTCCAGAGAGAGCTCAAGGGAGGCTGCTGGGCTGGGAGCCCTGCC 246  
 Db 578 GAGATGCTGTTTCCAGAGAGAGCTCAAGGGAGGCTGCTGGGCTGGGAGCCCTGCC 637  
 OY 247 AGGCTGTTGTCCTGTCAGAGGCTTTTGGGAGAGAGCTGCTAGAGCTGGAGGAGAC 306  
 Db 638 AGGCTGTTGTCCTGTCAGAGGCTTTTGGGAGAGAGCTGCTAGAGCTGGAGGAGAC 697  
 OY 307 TTCGGTGTGCAAGTTCGAGAGGCTGACAGTGAATCTGGGCGCAGGCTGAGCTGCTG 366  
 Db 698 TTCGGTGTGCAAGTTCGAGAGGCTGACAGTGAATCTGGGCGCAGGCTGAGCTGCTG 757  
 OY 367 GGTGAGAGAGAGCTGGAGCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 426  
 Db 758 GGTGAGAGAGAGCTGGAGCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 817  
 OY 427 GCATCTCTGCACTGGGATGGGGAGCCCTGTTAGGCGTGTACATATATCGGGGGCTGAA 486  
 Db 818 GCATCTCTGCACTGGGATGGGGAGCCCTGTTAGGCGTGTACATATATCGGGGGCTGAA 877  
 OY 487 CTCACCTTCAGCCCTTGGAGGAGGAGCCCTTAATCTGCTGGGGAGCTGGGGCTCAC 546  
 Db 878 CTCACCTTCAGCCCTTGGAGGAGGAGCCCTTAATCTGCTGGGGAGCTGGGGCTCAC 937  
 OY 547 ATCTACAGCCGGAAGAGTCTGCCAGGCGGTAAAGTCCATGTGCAAGTCAAGGCTCT 606  
 Db 938 ATCTACAGCCGGAAGAGTCTGCCAGGCGGTAAAGTCCATGTGCAAGTCAAGGCTCT 997  
 OY 607 CTTGGAAGCCCAAGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 666  
 Db 998 CTTGGAAGCCCAAGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1057  
 OY 667 CTGAGAGACACTGGTGGTGGAGATGACAGAGATGGCCGATTCACAGGTGGGGCTAAG 726  
 Db 1057 CTGAGAGACACTGGTGGTGGAGATGACAGAGATGGCCGATTCACAGGTGGGGCTAAG 1117

Db 1058 GTGAGACACTGGTGGTGGAGATGACAGAGATGGCCGATTCACAGGTGGGGCTAAG 1117  
 OY 727 CGTACCTGTTAAGATGATGGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGATCCG 786  
 Db 1118 CGTACCTGTTAAGATGATGGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGATCCG 1177  
 OY 787 AATCCTGTACCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 846  
 Db 1178 AATCCTGTACCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1237  
 OY 847 CAAGTGGGGCCCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 906  
 Db 1238 CAAGTGGGGCCCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1297  
 OY 907 AACACCCCTGAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 966  
 Db 1298 AACACCCCTGAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1357  
 OY 967 GACCTGTGTGAGTCTCCACTTGGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1026  
 Db 1358 GACCTGTGTGAGTCTCCACTTGGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1417  
 OY 1027 GACCCGCTGGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1086  
 Db 1418 GACCCGCTGGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1477  
 OY 1087 GCTCATGAGTGGGTCATGTCTTCAACATGGTCCATGCACTCAACGATCCATGATCAGT 1146  
 Db 1478 GCTCATGAGTGGGTCATGTCTTCAACATGGTCCATGCACTCAACGATCCATGATCAGT 1537  
 OY 1147 TTGAATGGGCTTTTGGAGACCTCTTCCCATGATGATGATGATGATGATGATGATGATGAT 1206  
 Db 1538 TTGAATGGGCTTTTGGAGACCTCTTCCCATGATGATGATGATGATGATGATGATGATGAT 1597  
 OY 1207 CCTGAGAGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1266  
 Db 1598 CCTGAGAGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1657  
 OY 1267 TATGGGACTGTCTTGGAGCAACAGAGGCTCCATGATGATGATGATGATGATGATGATGAT 1326  
 Db 1658 TATGGGACTGTCTTGGAGCAACAGAGGCTCCATGATGATGATGATGATGATGATGATGAT 1717  
 OY 1327 TACAAGAGAGAGATGACAAGGGGTAGTCTGA 1358  
 Db 1718 GGAAGAGACTATGATGTGACCGGCAAGTGCCA 1749  
 RESULT 3  
 AAX00725  
 ID AAX00725 standard; cDNA to mRNA; 4192 BP.  
 XX  
 AC AAX00725;  
 XX  
 DT 30-MAR-1999 (first entry)  
 XX  
 DE Human aggrecan degrading metalloproteinase 1 gene.  
 XX  
 KW Human; aggrecan degrading metalloproteinase; cartilage; proteoglycan;  
 KW interglobular domain; matrix metalloproteinase; bovine; interlucan-1beta;  
 KW primer; PCR; amplification; inhibitor; cleavage; inhibitor; ADMP; ss;  
 KW arthritis; joint injury; pseudogout.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 406..2919  
 FT /tag- a  
 FT /product- "ADMP-1"  
 FT /note- "aggrecan degrading metalloproteinase 1"  
 PN M09905291-A2.  
 XX  
 PD 04-FEB-1999.

XX 24-JUL-1998; 98WO-US15438.  
 XX 16-OCT-1997; 97US-0062169.  
 PR 25-JUL-1997; 97US-0053850.  
 PR 15-AUG-1997; 97US-0055836.  
 XX  
 PA (DUPO ) DU PONT PHARM CO.  
 PI Arner EC, Burn TC, Copeland RA, Decicco CP, Liu R, Magolda R;  
 PI Pratta M, Solomon KA, Tortorella MD, Tizaskos JM, Yang F;  
 DR WPI: 1999-142943/12.  
 P-PSDB: AAM75425.  
 PT New isolated aggrecan degrading metalloproteinases - used to develop  
 PT products for treating, e.g. osteoarthritis, joint injury, reactive  
 PT arthritis, psoriatic arthritis or juvenile rheumatoid arthritis  
 XX  
 PS Claim 12; Page 57-60; 73pp; English.  
 XX  
 CC This sequence represents the coding region for the human aggrecan  
 CC degrading metalloproteinase 1 (ADMP-1). ADMP-1 and ADMP-2 (AA00726) are  
 CC novel proteases that cleave the aggrecan (a major cartilage  
 CC (proteoglycan) between residues Glu373-Ala374 of the Interlobular domain  
 CC (compared with cleavage between Asn341-Phe342 by the matrix  
 CC metalloproteinases MMP-1, -2, -3, -7, -8, -9 and -13). ADMP-1 and ADMP-2  
 CC were isolated and purified from the conditioned media of bovine nasal  
 CC cartilage stimulated by interleukin-1beta. The purified proteases were  
 CC partially sequenced and primers were synthesised based on the resultant  
 CC amino acid sequences (AA00727-X00732 for ADMP-1 and AA00733-X00736 for  
 CC ADMP-2). These were used to isolate the corresponding genes from cDNA.  
 CC The ADMP polypeptides can be used for identifying inhibitors of ADMP  
 CC activity which would prevent cleavage of the aggrecan core protein,  
 CC thereby decreasing the loss of aggrecan from cartilage. Such inhibitors  
 CC can be used for treating diseases such as osteoarthritis, joint injury,  
 CC reactive arthritis, acute pyrophosphate arthritis (pseudogout), psoriatic  
 CC arthritis and juvenile rheumatoid arthritis.  
 CC  
 XX Sequence 4192 BP; 807 A; 1250 C; 1218 G; 917 T; 0 other:  
 SQ  
 Query Match 97.0%; Score 1318.4; DB 20; Length 4192;  
 Best Local Similarity 98.4%; Pred. No. 0;  
 Matches 1331; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

OY 427 GCATCTCTGCACTGGGATGGGAGGCCCTGTAGGCCGTTCATATATCGGGGGCTGAA 486  
 DB 823 GCATCTCTGCACTGGGATGGGAGGCCCTGTAGGCCGTTCATATATCGGGGGCTGAA 882  
 OY 487 CTCACCTTCACACCCCTGGAGGAGGACCCCTAACTCTGTGGGGAGCTGGGCTCAC 546  
 DB 883 CTCACCTTCACACCCCTGGAGGAGGACCCCTAACTCTGTGGGGAGCTGGGCTCAC 942  
 OY 547 ATCTAAGCCGGGAAGAGTCTGCCAGCGGTCAAGGCCCATGTGCAAGCTCAAGGCTCT 606  
 DB 943 ATCTAAGCCGGGAAGAGTCTGCCAGCGGTCAAGGCCCATGTGCAAGCTCAAGGCTCT 1002  
 OY 607 CTTGGAAGCCCGGAGCCCGGAGCCCGGAAGAGCCGCTTGTCTTCACTGATGATTT 666  
 DB 1003 CTTGGAAGCCCGGAGCCCGGAGCCCGGAAGAGCCGCTTGTCTTCACTGATGATTT 1062  
 OY 667 GTGGAGACCTGTGTGTGGCAGATGACAAGATGGCCGATTCACAGGTGGGGCTAAG 726  
 DB 1063 GTGGAGACCTGTGTGTGGCAGATGACAAGATGGCCGATTCACAGGTGGGGCTAAG 1122  
 OY 727 CGCTACCTGCTAAGAGTGTGGCAGAGGAGCCAGAGCCCTTCAAGACCCAGATCCGC 786  
 DB 1123 CGCTACCTGCTAAGAGTGTGGCAGAGGAGCCAGAGCCCTTCAAGACCCAGATCCGC 1182  
 OY 787 AATCTGTACGCTTGTGTGTGACTGGGCTAGTATCTGTGGGCTCAGGCGAGAGGGCCC 846  
 DB 1183 AATCTGTACGCTTGTGTGTGACTGGGCTAGTATCTGTGGGCTCAGGCGAGAGGGCCC 1242  
 OY 847 CAAGTGGGGCCAGAGTGTGGCAGAGCCCGGAGCTTGTGGCTGGAGGGGGCCCTC 906  
 DB 1243 CAAGTGGGGCCAGAGTGTGGCAGAGCCCGGAGCTTGTGGCTGGAGGGGGCCCTC 1302  
 OY 907 AACACCCCTGAGAGTGTGGAGCCCTGAGCACTTGTGACAGCCATCTGTTACCCGTCAG 966  
 DB 1303 AACACCCCTGAGAGTGTGGAGCCCTGAGCACTTGTGACAGCCATCTGTTACCCGTCAG 1362  
 OY 967 GACCTGTGTGGAGTGTCCACTTGGCAGACAGCGTGGATGTGATGTGGCAGCCGTCT 1026  
 DB 1363 GACCTGTGTGGAGTGTCCACTTGGCAGACAGCGTGGATGTGATGTGGCAGCCGTCT 1422  
 OY 1027 GACCCGGCTCGAGAGTGTGCCATTTGTGAGATGATGGGCTCCAGTCAAGCTTCACTGCT 1086  
 DB 1423 GACCCGGCTCGAGAGTGTGCCATTTGTGAGATGATGGGCTCCAGTCAAGCTTCACTGCT 1482  
 OY 1087 GCTCATGAACCTGGTATGTCTTCAACATGCTCATGACACACTCAAGCATGATCACT 1146  
 DB 1483 GCTCATGAACCTGGTATGTCTTCAACATGCTCATGACACACTCAAGCATGATCACT 1542  
 OY 1147 TTGAATGGGCTTTTGAAGCACTCTGCGCATGTGATGGCCCTGTATGTGATGTGAT 1206  
 DB 1543 TTGAATGGGCTTTTGAAGCACTCTGCGCATGTGATGGCCCTGTATGTGATGTGAT 1602  
 OY 1207 CCTGAGAGCCCTGTGTGCCCTCGAGTGGCCGCTTCATCACTGATGATGATGATGAT 1266  
 DB 1603 CCTGAGAGCCCTGTGTGCCCTCGAGTGGCCGCTTCATCACTGATGATGATGATGAT 1662  
 OY 1267 TATGGGCACTGTCTTTAGACAACCAAGAGGCTCAATTCATCTCTGTGATGGGGAC 1326  
 DB 1663 TATGGGCACTGTCTTTAGACAACCAAGAGGCTCAATTCATCTCTGTGATGGGGAC 1722  
 OY 1327 TACAAGAGCAGATGACACAAGGGGTAGTGCA 1358  
 DB 1723 GGAAGGACTATGATGTGACGCGCAAGTGCCA 1754  
 RESULT 4  
 ID AAS46100 standard; cDNA: 4406 BP.  
 XX AAS46100;  
 AC AAS46100;  
 XX  
 DT 18-DEC-2001 (first entry)  
 XX

DE Human DNA encoding PRO polypeptide sequence #176.  
XX  
XX PRO polypeptide: mammal; tumour; cancer; human; cattle; horse; sheep; ss;  
KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;  
KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;  
KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder;  
XX PCR primer.  
XX  
OS Homo sapiens.  
XX  
XX W0200168848-A2.  
XX  
XX 20-SEP-2001.  
XX  
XX 28-FEB-2001; 2001WO-US06520.  
XX  
XX 01-MAR-2000; 2000WO-US05601.  
XX 02-MAR-2000; 2000WO-US05841.  
XX 03-MAR-2000; 2000US-187202P.  
XX 06-MAR-2000; 2000US-186968P.  
XX 14-MAR-2000; 2000US-189320P.  
XX 14-MAR-2000; 2000US-189328P.  
XX 15-MAR-2000; 2000WO-US06884.  
XX 21-MAR-2000; 2000US-190828P.  
XX 21-MAR-2000; 2000US-191007P.  
XX 21-MAR-2000; 2000US-191048P.  
XX 21-MAR-2000; 2000US-191314P.  
XX 28-MAR-2000; 2000US-192655P.  
XX 29-MAR-2000; 2000US-193032P.  
XX 29-MAR-2000; 2000US-193053P.  
XX 30-MAR-2000; 2000WO-US08439.  
XX 04-APR-2000; 2000US-194449P.  
XX 11-APR-2000; 2000US-195975P.  
XX 11-APR-2000; 2000US-196000P.  
XX 11-APR-2000; 2000US-196187P.  
XX 11-APR-2000; 2000US-196900P.  
XX 11-APR-2000; 2000US-196820P.  
XX 18-APR-2000; 2000US-198121P.  
XX 18-APR-2000; 2000US-198585P.  
XX 25-APR-2000; 2000US-199397P.  
XX 25-APR-2000; 2000US-199550P.  
XX 25-APR-2000; 2000US-199634P.  
XX 03-MAY-2000; 2000US-201516P.  
XX 17-MAY-2000; 2000WO-US13705.  
XX 22-MAY-2000; 2000WO-US14042.  
XX 30-MAY-2000; 2000WO-US14941.  
XX 02-JUN-2000; 2000WO-US15266.  
XX 05-JUN-2000; 2000US-209832P.  
XX 28-JUL-2000; 2000WO-US20710.  
XX 22-AUG-2000; 2000US-064484P.  
XX 24-AUG-2000; 2000WO-US23328.  
XX 08-NOV-2000; 2000WO-US30952.  
XX 01-DEC-2000; 2000WO-US32678.  
XX 20-DEC-2000; 2000WO-US34956.  
XX  
XX (GETH ) GENENTECH INC.  
XX  
XX Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;  
XX Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;  
XX  
XX WPI: 2001-602746/68.  
XX P-PSDB: AAU29199.  
XX  
XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the  
XX presence of tumours, such as prostate and breast tumours, in mammals and  
XX to screen for modulators of the compounds -  
XX  
XX Claim 2: Fig 351, 774pp; English.  
XX  
XX Sequences AAS45925-AAS46231 represent DNA molecules encoding and PCR  
XX primers for PRO polypeptides of the invention. The sequences of the  
XX invention can be used to detect the presence of a tumour in a mammal by

CC comparing the level of expression of a PRO polypeptide in a test sample  
CC of cells from the animal and a control sample indicates the presence of a  
CC higher level of expression in the test sample indicates the presence of a  
CC tumour in the mammal. Mammals include dogs, cats, cattle, horses, sheep,  
CC pigs, goats and rabbits but are preferably human. The polypeptides can be  
CC used to stimulate tumour necrosis factor (TNF) alpha release from human  
CC blood, when contacted with it. A specific polypeptide can be used to  
CC stimulate the proliferation or differentiation of chondrocyte cells. The  
CC PRO proteins can be used to determine the presence of tumours and also  
CC susceptibility to tumour development, particularly adrenal, lung, colon,  
CC breast, prostate, rectal, cervical, or liver tumours. In mammalian  
CC subjects. The oligonucleotide probes specific for the PRO nucleic acids  
CC can be used for genetic analysis of individuals with genetic disorders.  
XX  
XX Sequence 4406 BP: 931 A; 1275 C; 1241 G; 959 T; 0 other:  
XX  
XX Query Match 96.9%; Score 1316.8; DB 22; Length 4406;  
XX Best Local Similarity 98.4%; Pred. No. 0;  
XX Matches 1330; Conservative 0; Mismatches 22; Indels 0; Gaps 0;  
OY 7 GCCATGTCACAGACAGGCGCATCCGGAGGGGCTTGGCAGGGGCGTGGCTGTGGGA 66  
DB 416 GCCATGTCACAGACAGGCGCATCCGGAGGGGCTTGGCAGGGGCGTGGGGA 475  
OY 67 GCCCAACCCCTGCTCTCTCCATTTGTGCGCTCTCTGCTGGTGGCTGCTTCTG 126  
DB 476 GCCCAACCCCTGCTCTCTCTCCATTTGTGCGCTCTCTGCTGGTGGCTTCTG 535  
OY 127 CTACTGTGGCT 186  
DB 536 CTACTGTGGCT 595  
OY 187 GAGATCGTGTTCAGAGAGAGCTCAAGGGGAGGCTGCGCTGGGAGCCCTGCTCC 246  
DB 596 GAGATCGTGTTCAGAGAGAGCTCAAGGGGAGGCTGCGCTGGGAGCCCTGCTCC 655  
OY 247 AGGCTGTTGTGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 306  
DB 656 AGGCTGTTGTGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 715  
OY 307 TCCGGTGTGACAGGTGAGGGGCTGACAGTACCTGAGGGGAGGCGCTGAGTCTG 366  
DB 716 TCCGGTGTGACAGGTGAGGGGCTGACAGTACCTGAGGGGAGGCGCTGAGTCTG 775  
OY 367 GGTGAGAGAGAGCTGAGACCTGACTGACCTGACCTCAATGAGAGTCCGGAGTGG 426  
DB 776 GGTGAGAGAGAGCTGAGACCTGACTGACCTGACCTCAATGAGAGTCCGGAGTGG 835  
OY 427 GCATCTCTGACCTGAGATGGGAGCCCTGTTAGGCGTTCATATTCGGGGGCTGAA 486  
DB 836 GCATCTCTGACCTGAGATGGGAGCCCTGTTAGGCGTTCATATTCGGGGGCTGAA 895  
OY 487 CTCACACTTCAGCCCTGAGGAGGAGGAGCCCTGACTGCTGGGGGAGCTGGGCTTAC 546  
DB 896 CTCACACTTCAGCCCTGAGGAGGAGGAGCCCTGACTGCTGGGGGAGCTGGGCTTAC 955  
OY 547 ATCTACGCGGAGAGAGTCTGCGAGCGTCAAGTCTCCATGTGCAAGTCAAGCTCT 606  
DB 956 ATCTACGCGGAGAGAGTCTGCGAGCGTCAAGTCTCCATGTGCAAGTCAAGCTCT 1015  
OY 607 CTTGGAAGCCCGAGCCGAGAGAGCCGTTGCTTCACTGAGTGAATTT 666  
DB 1016 CTTGGAAGCCCGAGCCGAGAGAGCCGTTGCTTCACTGAGTGAATTT 1075  
OY 667 GTGGAGACACTGGTGTGGAGATGACAAAGTGGCGATTCACAGGTGGGGGCTAAG 726  
DB 1076 GTGGAGACACTGGTGTGGAGATGACAAAGTGGCGATTCACAGGTGGGGGCTAAG 1135  
OY 727 CGCTACCTGCTAACAGTGTGACAGAGCCGCTTCAACACCAAGCATCCG 786  
DB 1136 CGCTACCTGCTAACAGTGTGACAGAGCCGCTTCAACACCAAGCATCCG 1195  
OY 787 AATCTGTACGTTGTGTGACTGCGCTAGTGTCTGGGCTCAGGCGAGAGGGGCC 846

Db 1196 AATCCGTCAAGCTTGGTGGAGTCCGCTAGTCATCTGGGGTCAGGCGAGGGGCCCC 1255  
OY 847 CAAGTGGGGCCAGTGTGCTGGCCAGACCTGGCCGAGCTTGTGTCGGGAGGGGGCTTC 906  
Db 1256 CAAGTGGGGCCAGTGTGCTGGCCAGACCTGGCCGAGCTTGTGTCGGGAGGGGGCTTC 1315  
OY 907 AACACCCCTGAGAGTCTGGAGCCCTGACCACTTTGACACAGCCATTCCTGTTTACCCTCAG 966  
Db 1316 AACACCCCTGAGAGTCTGGAGCCCTGACCACTTTGACACAGCCATTCCTGTTTACCCTCAG 1375  
OY 967 GACCTGTGTGAGATCTTCACATTGCGACAGCTGGGTATGGCTATGTGGGACCGTCTGT 1026  
Db 1376 GACCTGTGTGAGATCTTCACATTGCGACAGCTGGGTATGGCTATGTGGGACCGTCTGT 1435  
OY 1027 GACCCGGCTGGAGTCTGTGCTGATTTGGAGATGATGGCTCCAGTCCAGCTTCACTGCT 1086  
Db 1436 GACCCGGCTGGAGTCTGTGCTGATTTGGAGATGATGGCTCCAGTCCAGCTTCACTGCT 1495  
OY 1087 GCTCATGAATGGGTGATGTCTTCAACATGCTCCATGACAACTCAAGCCATGATCACT 1146  
Db 1496 GCTCATGAATGGGTGATGTCTTCAACATGCTCCATGACAACTCAAGCCATGATCACT 1555  
OY 1147 TTGAATGGGCTTTGAGCACTCTGCGCATGTGATGGCCCTGTGATGGCTCATGTGAT 1206  
Db 1556 TTGAATGGGCTTTGAGCACTCTGCGCATGTGATGGCCCTGTGATGGCTCATGTGAT 1615  
OY 1207 CCTGAGAGAGCCCTGATCCCTGCTGACAGTGGCCGCTTCATGACTGCTTCTGCAATGAG 1266  
Db 1616 CCTGAGAGAGCCCTGATCCCTGCTGACAGTGGCCGCTTCATGACTGCTTCTGCAATGAG 1675  
OY 1267 TATGGGCACTGTCTTACACAACAGAGAGCTCCATTTGCATCTGCTGTGATGGGGAC 1326  
Db 1676 TATGGGCACTGTCTTACACAACAGAGAGCTCCATTTGCATCTGCTGTGATGGGGAC 1735  
OY 1327 TACAGAGACAGCATGACAAGGGGTAGTGCA 1358  
Db 1736 GCGAAGGACTATGATCTGACCCGCCAGTGCCA 1767

RESULT 5  
AAA37111  
ID AAA37111 standard; cDNA; 4407 BP.  
XX  
AC AAA37111;  
XX  
DT 08-AUG-2000 (first entry)  
XX  
DE Human PRO1563 (UNQ769) cDNA sequence SEQ ID NO:316.  
KW Human; PRO polypeptide; membrane bound protein; receptor; diagnosis;  
KW Transmembrane; secretion; Immunoadhesion; pharmaceutical; screening;  
KW ss.  
OS Homo sapiens.  
XX  
PN WO200012708-A2.  
XX  
PD 09-MAR-2000.  
XX  
PF 01-SEP-1999; 99WO-US20111.  
XX  
PR 01-SEP-1998; 98US-0098716.  
PR 01-SEP-1998; 98US-0098749.  
PR 01-SEP-1998; 98US-0098750.  
PR 02-SEP-1998; 98US-0098803.  
PR 02-SEP-1998; 98US-0098821.  
PR 02-SEP-1998; 98US-0098843.  
PR 09-SEP-1998; 98US-0099536.  
PR 09-SEP-1998; 98US-0099596.  
PR 09-SEP-1998; 98US-0099598.  
PR 09-SEP-1998; 98US-0099602.  
PR 09-SEP-1998; 98US-0099642.

PR 10-SEP-1998; 98US-0099741.  
PR 10-SEP-1998; 98US-0099754.  
PR 10-SEP-1998; 98US-0099763.  
PR 10-SEP-1998; 98US-0099792.  
PR 10-SEP-1998; 98US-0099808.  
PR 10-SEP-1998; 98US-0099812.  
PR 10-SEP-1998; 98US-0099815.  
PR 10-SEP-1998; 98US-0099816.  
PR 15-SEP-1998; 98US-0100385.  
PR 15-SEP-1998; 98US-0100386.  
PR 15-SEP-1998; 98US-0100390.  
PR 16-SEP-1998; 98US-0100584.  
PR 16-SEP-1998; 98US-0100627.  
PR 16-SEP-1998; 98US-0100661.  
PR 16-SEP-1998; 98US-0100662.  
PR 16-SEP-1998; 98US-0100664.  
PR 17-SEP-1998; 98US-0100683.  
PR 17-SEP-1998; 98US-0100684.  
PR 17-SEP-1998; 98US-0100710.  
PR 17-SEP-1998; 98US-0100711.  
PR 17-SEP-1998; 98US-0100919.  
PR 17-SEP-1998; 98US-0100930.  
PR 18-SEP-1998; 98US-0100848.  
PR 18-SEP-1998; 98US-0100849.  
PR 18-SEP-1998; 98US-0101014.  
PR 18-SEP-1998; 98US-0101068.  
PR 18-SEP-1998; 98US-0101071.  
PR 22-SEP-1998; 98US-0101279.  
PR 23-SEP-1998; 98US-0101471.  
PR 23-SEP-1998; 98US-0101472.  
PR 23-SEP-1998; 98US-0101474.  
PR 23-SEP-1998; 98US-0101475.  
PR 23-SEP-1998; 98US-0101476.  
PR 23-SEP-1998; 98US-0101477.  
PR 23-SEP-1998; 98US-0101479.  
PR 24-SEP-1998; 98US-0101738.  
PR 24-SEP-1998; 98US-0101741.  
PR 24-SEP-1998; 98US-0101743.  
PR 24-SEP-1998; 98US-0101915.  
PR 24-SEP-1998; 98US-0101916.  
PR 29-SEP-1998; 98US-0102207.  
PR 29-SEP-1998; 98US-0102240.  
PR 29-SEP-1998; 98US-0102307.  
PR 29-SEP-1998; 98US-0102307.  
PR 29-SEP-1998; 98US-0102307.  
PR 29-SEP-1998; 98US-0102330.  
PR 29-SEP-1998; 98US-0102331.  
PR 30-SEP-1998; 98US-0102484.  
PR 30-SEP-1998; 98US-0102487.  
PR 30-SEP-1998; 98US-0102570.  
PR 30-SEP-1998; 98US-0102571.  
PR 01-OCT-1998; 98US-0102684.  
PR 01-OCT-1998; 98US-0102687.  
PR 02-OCT-1998; 98US-0102965.  
PR 06-OCT-1998; 98US-0103258.  
PR 06-OCT-1998; 98US-0103449.  
PR 07-OCT-1998; 98US-0103314.  
PR 07-OCT-1998; 98US-0103315.  
PR 07-OCT-1998; 98US-0103315.  
PR 07-OCT-1998; 98US-0103328.  
PR 07-OCT-1998; 98US-0103395.  
PR 07-OCT-1998; 98US-0103396.  
PR 07-OCT-1998; 98US-0103401.  
PR 08-OCT-1998; 98US-0103633.  
PR 08-OCT-1998; 98US-0103678.  
PR 08-OCT-1998; 98US-0103679.  
PR 08-OCT-1998; 98US-0103711.  
PR 14-OCT-1998; 98US-0104257.  
PR 20-OCT-1998; 98US-0104987.  
PR 20-OCT-1998; 98US-0105000.  
PR 20-OCT-1998; 98US-0105002.  
PR 21-OCT-1998; 98US-0105104.  
PR 22-OCT-1998; 98US-0105169.  
PR 22-OCT-1998; 98US-0105266.  
PR 26-OCT-1998; 98US-0105693.  
PR 26-OCT-1998; 98US-0105694.



Db 1556 TTGAATGGCCCTTTGAGCACTCTCCCATGTCATGATGCGCCCTGTGTGATGCTCATGTGAT 1615  
OY 1207 CCTGAGAGCCCTGTGCTCCCTGTGACAGTCCCGCTTTCATGACTGACTTCTCGACAAATGAGC 1266  
Db 1616 CCGAGAGAGCCCTGTGCTCCCTGTGACAGTCCCGCTTTCATGACTGACTTCTCGACAAATGAGC 1675  
OY 1267 TATGGCACTGTCTCTTACGACAAACAGAGGCTTCATGCTGCTGTGACTGAGGAGC 1326  
Db 1676 TATGGCACTGTCTCTTACGACAAACAGAGGCTTCATGCTGCTGTGACTGAGGAGC 1735  
OY 1327 TACAGAGAGAGAGTGAACAGAGGCTGTGCTGCA 1358  
Db 1736 GCGAAGAGAGTGTGATGCTGACGCGCAGTGCCA 1767  
RESULT 6  
AAF54426  
ID AAF54426 standard; DNA; 4407 BP.  
XX AAF54426;  
AC AAF54426;  
XX 02-APR-2001 (first entry)  
DT 02-APR-2001 (first entry)  
XX Probe #46 used in the identification of proteins.  
DE Secreted; transmembrane; gene therapy; ss.  
XX Unidentified.  
OS Unidentified.  
XX MO200078961-A1.  
XX 28-DEC-2000.  
PD 28-DEC-2000.  
XX 18-FEB-2000; 2000MO-US04342.  
PF 18-FEB-2000; 2000MO-US04342.  
XX 23-JUN-1999; 99US-0141037.  
PR 20-JUL-1999; 99US-0144758.  
PR 26-JUL-1999; 99US-0145698.  
PR 01-SEP-1999; 99MO-US20111.  
PR 29-OCT-1999; 99US-0162506.  
PR 30-NOV-1999; 99MO-US28313.  
PR 02-DEC-1999; 99MO-US28551.  
PR 16-DEC-1999; 99MO-US30095.  
PR 05-JAN-2000; 2000MO-US00219.  
PR 06-JAN-2000; 2000MO-US00376.  
XX (GENTH ) GENE/TECH INC.  
PA Baker KP, Botstein D, Desnovers L, Eaton DL, Ferrara N, Fong S;  
XX Gao W, Goddard A, Godowski PJ, Grimaldi CJ, Gutney AL, Hillan KJ;  
PI Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D;  
PI Watanabe CK, Williams PM, Wood WI;  
XX MPI: 2001-071395/08.  
DR MPI: 2001-071395/08.  
PT Secreted and transmembrane proteins and nucleic acids designated PRO,  
PT useful as hybridization probes. In chromosome and gene mapping and gene  
therapy -  
XX Example 91: Page 463; 787pp: English.  
XX The present invention relates to secreted and transmembrane proteins.  
XX These proteins and the DNA encoding them may be used as hybridization  
XX probes, in chromosome and gene mapping and in the generation of  
XX anti-sense RNA and DNA. They may also be used to generate either  
XX transgenic animals or knockout animals which are in turn useful for  
XX development and screening of therapeutically useful reagents.  
XX The nucleic acids may also be used in gene therapy.  
SO Sequence 4407 BP; 932 A; 1275 C; 1241 G; 959 T; 0 other;  
Query Match 96.9%: Score 1316.8; DB 22; Length 4407;  
Best Local Similarity 98.4%: Pred. No. 0;

Matches 1330; Conservative 0; Mismatches 22; Indels 0; Gaps 0;  
OY 7 GCATGTCACAGACAGAGTCCGATCCGGAGAGGCTTGGACAGGCGCTGTGAGGA 66  
Db 416 GCATGTCACAGACAGAGTCCGATCCGGAGAGGCTTGGACAGGCGCTGTGAGGA 475  
OY 67 GCCAAGCCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 126  
Db 476 GCCAAGCCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 535  
OY 127 CTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 186  
Db 536 CTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 595  
OY 187 GAGATGCTGCTTTCAGAGAGCTCAACAGGAGAGCTGCTGCTGCTGCTGCTGCTG 246  
Db 596 GAGATGCTGCTTTCAGAGAGCTCAACAGGAGAGCTGCTGCTGCTGCTGCTGCTG 655  
OY 247 AGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 306  
Db 656 AGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 715  
OY 307 TCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 366  
Db 716 TCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 775  
OY 367 GGTGAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 426  
Db 776 GGTGAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 835  
OY 427 GCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 486  
Db 836 GCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 895  
OY 487 CTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 546  
Db 896 CTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 955  
OY 547 ATCTAAGAGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 606  
Db 956 ATCTAAGAGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1015  
OY 607 CTGTAAG 666  
Db 1016 CTGTAAG 1075  
OY 667 GTGAG 726  
Db 1076 GTGAG 1135  
OY 727 CGCTAAG 786  
Db 1136 CGCTAAG 1195  
OY 787 AATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 846  
Db 1196 AATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1255  
OY 847 CAAGTGGAG 906  
Db 1256 CAAGTGGAG 1315  
OY 907 AACAGCCCTGAG 966  
Db 1316 AACAGCCCTGAG 1375  
OY 967 GACCTGCTGAG 1026  
Db 1376 GACCTGCTGAG 1435  
OY 1027 GACCGGCTGAG 1086  
Db 1436 GACCGGCTGAG 1495



[illegible]

Query Match	20.7%	Score 281;	DB 20;	Length 4712;
Best Local Similarity	62.2%	Pred. No. 1.9e-59;		
Matches 456;	Conservative 2;	Mismatches 272;	Indels 3;	Gaps 1

Oy 626 GACCCCGAAGAGCCAAAGCGCTTGGCTTCACATGAGTAGATTGTGGAAACACTGTGGTGCG 685  
||| ||| ||| ||| | ||| ||| ||| ||| |||  
Db 1244 GAAGCATTAAGAAGAACGCATTTGTCTCCAGTCACC GCCTATGTGGAAACCAATCGTTGTGG 1303

OY	686	CAGATGACAGATATGGCCGCAATTCACAGGTCGCGGGCTAAAGCGCTACCTGCTACAGATGA	745
Db	1304	CAGACGATCGATATGGCAGAAATTCACAGGCGAGTGGTCTAAAGCATTTACCTTCCACGTTGT	1365
OY	746	TGGCAGCAGCAGCCACGAGCCCTTCACAGCACCACAGCATCCGCAATCCTGTACACTGGTGG	805
Db	1364	TTTTCGGTGGCAGGCCAGATTTGTCMAAACACCCGASATTTGGTAAATTCAGTTACCTCGTGG	1422
OY	806	TGACTCGGCTAGTATCTCTGGGGTCAGGCGAGAGAGGGGCCCAATGGGGCCCACTGCTG	865
Db	1424	TGTGTAAATCTTGGTATCTACACGATGAACAGAAAGGGGGGGAAGTACCTTCATGTGTG	1483
OY	866	CCGACAGCCCTGCGCAGCTTCTGTGCTGGCAGCGGGGGCCCTCAACACCCCTGAGAGCTCG	925
Db	1484	CCCTACGCTCGGGAACCTTTTGGCACTGGCAGAAAGCACAACCCACCACTGATGACGGG	1543
OY	926	ACCTGACCACTTTTGACACAGCCATTCTGTTTACCCTGACGACCTGTGTGGATGTCCA	985
Db	1544	ATGAGAGCACTATAGACACAGCAATTTCTTTCACGAGACAGGACTTGTGTGGGTCCAGA	1603
OY	986	CTTGCGACACCGTGGTATGATGGCTGATGTGGGCGCCGCTGTGACCCGGCTGGAGCTGTG	1045
Db	1604	CATGTGATACCTCTGGGATGGCTGATGTGGATCTGTGTGATCCGACAGAACTGCT	1666
OY	1046	CCATTGTGGAGGATGATGGGCTCCAGTCAGCCCTTCACTGCTGTCATGAATGAGTGGTCATG	1105
Db	1664	CCGTCATAGAAAGATGATGGTTTCAAGAGTCGCTTCCACACACCCCATGATTAAGGCCACG	1723
OY	1106	TCTTCACATGCTGTCATGACAACTCCGACAGCTGATCGTTGTAATGGGCGCTTGAGACA	1165
Db	1724	TGTTTAAATATCCCACTATGATGATGACAAAGCAGTGTCCGACCTTAATGG--TGTAAAC	1780
OY	1166	CCCTCGCCATGTCATGAGCCCTGTGATGAGCTCATGTGATCCTGAGAGGCCCTGCTCC	1225
Db	1781	AGATTTCCCACTATGATGGCTCAATCTTCCAAACCTGGACACAGCCAGCTTGCTC	1848
OY	1226	CCCTGACATGGCCGTTTATATACGACTTCTCTGGACATGGCTATAGGCGCATGTCTCTTAG	1285
Db	1841	CTTCAGATGGCTCATGATGATTAATCAATTTCTGGATATATGTCATGGGAAATGTTTGAATGG	1900
OY	1286	ACAACACAGAGGCTCATTTGCACTATGCTCCTGTGACTGGGACTACAGAGCAGATGACA	1345
Db	1901	ACAAGCCTCAAGAAATCCCATACAGCTCCACAGGCAATCTCCCTGGCACTGTGATGATGCCA	1960
OY	1346	AGGGGTAGTGTGA 1358	
Db	1961	ACCGGACAGTGCCA 1973	

RESULT 9	
AAH20224	
ID	AAH20224 standard; cDNA; 2653 BP.
XX	
AC	AAH20224;
XX	
DT	07-AUG-2001 (first entry)
XX	
DE	Human ADAM-type metalloprotease MDTs4-encoding cDNA, SEQ ID NO:5
XX	
KW	Human; MDTs4; ADAM-type metalloprotease; drug screening;
XX	
OS	A Disintegrin And Metalloprotease; cancer; arthritis; ss.
XX	
XX	Homo sapiens.
XX	
FT	Location/Qualifiers
FT	1..2853
CDS	/tag= a /product= "Human MDTs4" /note= "ADAM-type metalloprotease"
XX	
FT	
FT	
PN	JP2001017183-A.
XX	
XX	
PD	23-JAN-2001.



```

XX 09-JUL-1999; 99JP-0196584.
PF ACAACAGAGGCTCCATGATGCTGCTGAGTGGGAGTCAAGAGCATGATGACA 1345
PR 09-JUL-1999; 99JP-0196584.
XX (YAMA ) YAMANOUCHI PHARM CO LTD.
PA WPI: 2001-275950/29.
XX P-PSDB: AAB73549.
DR A new metal protease and its preparation for use as an anti-cancer and
XX anti-arthritis therapeutic.
PT Example 1; Page 14-15; 22pp; Japanese.
XX
XX The invention relates to the novel human ADAM (A Disintegrin And
XX Metalloprotease)-type metalloproteases MDR54 (AAB73549) and MDR55
XX (AAB73550). The metalloproteases can be used for the treatment of
XX cancers and arthritis. The invention also relates to the genes encoding
XX MDR54 and MDR55, vectors and host cells containing the MDR54 or MDR55
XX genes, the recombinant production of MDR54 and MDR55, and antibody
XX specific for MDR54 or MDR55, and methods of screening for compounds
XX which modulate the activity of MDR54 and/or MDR55. The present
XX sequence represents cDNA encoding human MDR54.
SQ Sequence 2853 BP; 704 A; 734 C; 816 G; 599 T; 0 other;

Query Match 20.6%; Score 280.2; DB 22; Length 2853;
Best Local Similarity 62.3%; Pred No. 2,6e-59;
Matches 457; Conservative 0; Mismatches 273; Indels 3; Gaps 1;

QY 636 GACCCGGAAGAGCCGCTTGTCTCAGTAGATTTGTGGAGACATGCTGTGG 685
DB 686 GAAGCATTAAGAAAGACATTTGTCTCAGTAGATTTGTGGAGACATGCTGTGG 745
QY 686 CAGATGACAAAGATGCGCCGATTCACGCGGGGCTAAAGCGCTACCTGCTAACAGTGA 745
DB 746 CAGACCACTGCGAGATTCACGCGGAGTGTCTAAAGCATTAACCTTCACAGCTGT 805
QY 746 TGGCAGACGACCCCAAGGCTTCAAGACACCAAGCATTCCTGTCAGCTTGTGG 805
DB 806 TTTGCGTGGACACCGCATTTGTACAAACACCCGACATTCGTATTAAGTACCTGTGG 865
QY 806 TGACTGCGCTAGTATCCTGCGGCTGAGCGGAGAGGGGCCCAAGTGGGGCCAGTCTG 865
DB 866 TGGTGAAGATCTTGGTATCCACGATGAACAGAGGGGCCGAGACCTCCAAATGCTG 925
QY 866 CCCAGACCTGGCGAGCTTCTGCTGGCAGACGGGGGCTCAACACCCCTGAGAGCTGG 925
DB 926 CCGTCACTCTGGGAACTTTTGCACAGACAGACGACACCCACCACTGACCGGG 985
QY 926 ACCCTGACCACTTTGACAGACGATCTGTATACCGGTGAGACCTGTGGAGTCTCA 985
DB 986 ATGCAGAGACATGATGACACAGCAATTTCTTTACCAAGACAGACACTTTGGTCCCA 1045
QY 986 CTTCGACACAGCTGGGTATGCTGATGTGGGACACCGCTGTGACCCGGCTGGAGCTGTG 1045
DB 1046 CATGTGATGATCTTGTGGGATGCTGTGATGTGGGATGCTGTGATGCTGTGATGCTGTG 1105
QY 1046 CATGTGATGATGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTG 1105
DB 1106 CCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1165
QY 1106 TCTTCAACATGCTCCATGACAACTCCAAAGCCATGATGATGATGATGATGATGATGATGATG 1165
DB 1166 TCTTCAACATGCTCCATGACAACTCCAAAGCCATGATGATGATGATGATGATGATGATGATG 1222
QY 1166 CCTCTGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1225
DB 1223 AGGATTTCCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1282
QY 1226 CCGGAGTCCCGGCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1285

```

```

DB 1283 CTTCGATGCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1342
QY 1286 ACAACAGAGGCTCCATGATGCTGCTGAGTGGGAGTCAAGAGCATGATGACA 1345
DB 1343 ACAAGCTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1402
QY 1346 AGGCGTGGCTGCA 1358
DB 1403 ACCGGCAGTGCCA 1415

RESULT 10
AAZ32000
ID AAZ32000 standard; cDNA; 3261 BP.
XX
XX AAZ32000;
AC 10-JAN-2000 (first entry)
XX
XX Human METH1 encoding cDNA.
DE
XX
XX Human: METH1: anti-angiogenic; metalloprotease thrombospondin;
XX cancer; diagnosis; hyperproliferative disorder; autoimmune disease;
XX angiogenesis inhibitor; abnormal wound healing; inflammation;
XX rheumatoid arthritis; psoriasis; endometrial bleeding disorder;
XX diabetic retinopathy; macula degeneration; haemangioma; detection;
XX arterial-venous malformation; immune deficiency; ss.
XX
XX Homo sapiens.
XX
XX W09937660-A1.
XX
XX 29-JUL-1999.
XX
XX 22-JAN-1999; 99WO-US01313.
XX
XX 23-JAN-1998; 98US-0072298.
XX
XX 28-AUG-1998; 98US-0098539.
XX
XX (IRUE/) IRUELA-ARISPE L.
XX (HAST/) HASTINGS G A.
XX (RUBE/) RUBEN S M.
XX
XX IrueLA-Arispe L, Hastings GA, Ruben SM.
XX
XX WPI: 1999-590684/50.
XX
XX P-PSDB: AAY49501.
XX
XX New isolated metalloprotease thrombospondin polypeptides, useful for
XX treating hyperproliferative disorders, cancers or autoimmune disorders
XX
XX Claim 3; Fig 1; 457pp; English.
XX
XX AAZ32000 and AAZ32001 encode, and AAY49501 and AAY49502 represent, human
XX metalloprotease thrombospondin (METH) proteins METH1 and METH2
XX respectively. METH1 and METH2 have been found to be potent inhibitors of
XX angiogenesis both in vitro and in vivo. They can be used for treating
XX cancer and other disorders related to angiogenesis including abnormal
XX wound healing, inflammation, rheumatoid arthritis, psoriasis,
XX endometrial bleeding disorders, diabetic retinopathy, some forms of
XX macula degeneration, haemangiomas, and arterial-venous malformations.
XX They may be useful in treating deficiencies or disorders of the immune
XX system, by activating or inhibiting the proliferation, differentiation,
XX or mobilisation (chemotaxis) of immune cells. The etiology of these
XX immune deficiencies or disorders may be genetic, somatic, such as
XX cancer or some autoimmune disorders, acquired (e.g. by chemotherapy or
XX toxins), or infectious. They can also be used to treat inflammatory
XX conditions, both chronic and acute conditions. The products can also be
XX used for detection and diagnosis. AAZ32002 to AAZ32080, and AAY49503 to
XX AAY49511 represent sequences given in the exemplification of the present
XX invention.

```

Sequence 3261 BP; 844 A; 789 C; 924 G; 700 T; 4 other;

Query Match 20.6%: Score 280.2; DB 20; Length 3261;

Best Local Similarity 62.3%: Pred. No. 2.7e-59; Matches 457; Conservative 0; Mismatches 273; Indels 3; Gaps 1;

OY 626 GACCCGGAAGAGCCGCTTTGCTTCACTAGTAGATTGTTGGAGACAGCTGGTG 685  
 DB 686 GAAGCATAGAAAGAGGATTTGTCCAGTCACCGCTATGTGGAACCATGCTTGTG 745  
 OY 686 CAGATGACAAAGATGGCCGATTCACAGGTGGGGGCTAAAGGCTACCTGCTAACAGTGA 745  
 DB 746 CAGACCAAGTCGAGGATTCACAGGAGTGTCTTAAGCATTAACCTTCACCTTGT 805  
 OY 746 TGGCAGACAGCAGCCAGGCTTCAAGCAGCAGCATCCGCAATCTCTCAGCTTGGTG 805  
 DB 806 TTTCCGTGGCAGCAGCATTTGTACAAACACCCAGCATTTGTAATTCAGTTAGCTGGTG 865  
 OY 806 TGAATCGGCTAGTATCTCTGGGGTACGGCAGAGAGGGGCCCAAGTGGGCCAGTGTG 865  
 DB 866 TGGTCAAGATCTTGTGATCAACAGATGAACAGAGGGGCCGGAAGTCACTCAATGTG 925  
 OY 866 CCCAGACCCCTGGGAGCTTGTGCTGGCAGGGGGCTCAACACCCCTGAGGAGTGG 925  
 DB 926 CCTCTACTGTGGGAGACTTTTGCAGCTGGCAGAGACAGACACACCCAGTACCGGG 985  
 OY 926 ACCCTGACCACTTTGACAGAGCATCTGTTTACCCGTCAGAGACCTGTGTGAGTCTCA 985  
 DB 986 ATGCAGAGCATATGACAGAGCAATCTTTTACAGAGAGCATTTGTGGTGCCACA 1045  
 OY 986 CTTCGACAGCCTGGGTATGGCTGATGTGGCAGCCCTGTGTGACCCGGCTCGAGCTGTG 1045  
 DB 1046 CAGTGTATCTTGGAGTGGCTGATGTGGAACCTGTGTGATCCGACAGAAAGCTGCT 1105  
 OY 1046 CCATTGTGAGATGATGGGCTCCAGCTCAGCTTCTCTCTCTATGTAAGTGGTCTATG 1105  
 DB 1106 CCGTCTATGAAGATGATGTTTACAAAGCTGCCCTTCACACAGCCATGATTAAGGCA 1165  
 OY 1106 TCTTCAACATGCTCCATGACAACTCCAGCCATGATCACTGTTGAATGGGCTTTGAGA 1165  
 DB 1166 TCTTTTAACTGCCACATGATGATGAACAGCTGTGCCAGCTTAATG---TGTGAAC 1222  
 OY 1166 CCTGTGGCATGATGAGGGCCCTGTGATGGCTCATGTGATCTGTAGAGCCCTGCTCC 1225  
 DB 1223 AGCATTTCCCAATGATGGGCTCAATGCTTTTCAACCTGACACAGCCAGCTTGTCTC 1282  
 OY 1226 CCTGACGTCGCGCTTCACTACATCTCTGACAAATGGATATGGGCACTGCTTTAG 1285  
 DB 1283 CTTCGATGCTTCAATGATTAATCATCTTTCTGATTAATGCTGAGGAAATGTTGATG 1342  
 OY 1286 ACAAAACAGAGGCTTCATTTGATCTGCTGCTGATGGGACCTTACAGAGCATGACA 1345  
 DB 1343 ACAAGCTCAGATGCCATACAGCTCCAGGGAGCTCCCTGGCAGCTGCTAGATGCCA 1402  
 OY 1346 AGCGGTAGTCCA 1358  
 DB 1403 ACCGGCAGTGCCA 1415  
 RESULT 11  
 AAC90057  
 ID AAC90057 standard: DNA: 3261 BP.  
 AC AAC90057;  
 XX  
 XX  
 DT 19-MAR-2001 (first entry)  
 XX  
 XX Human METH1 coding sequence.  
 DE  
 XX Human METH1 coding sequence.  
 KW Human: METH1; metalloproteinase; thrombospondin; angiogenesis inhibition;  
 KW cancer therapy; benign tumor; ocular angiogenic disease;  
 KW rheumatoid arthritis; psoriasis; wound healing; endometriosis;  
 KW vasculogenesis; granulation; hypertrophic scar; nonunion fracture;

KW scleroderma, trachoma; vascular adhesion; myocardial angiogenesis;  
 KW coronary collateral; cerebral collateral; arteriovenous malformation;  
 KW ischemic limb angiogenesis; Osler-Webber syndrome; chromosome 21q21;  
 KW plaque neovascularisation; telangiectasia; haemophilic joint;  
 KW angiodysplasia; fibromuscular dysplasia; wound granulation;  
 KW Crohn's disease; atherosclerosis; birth control; ss.

Homo sapiens.  
 MO200071577-A1.

30-NOV-2000.

25-MAY-2000; 2000MO-US14462.

25-MAY-1999; 99US-0318208.

20-JUL-1999; 99US-0144882.

10-AUG-1999; 99US-0147823.

13-AUG-1999; 99US-0373658.

22-DEC-1999; 99US-0171503.

22-FEB-2000; 2000US-0183792.

(HUMA-) HUMAN GENOME SCI INC.

(SMK-) SMITHKLINE BEECHAM CORP.

(BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

(IRUE-) IRUELA-ARISPE L.

(HAST-) HASTINGS G A.

(RUBE-) RUBEN S M.

(JONA-) JONAK Z L.

(TRUL-) TRULLI S H.

(FORN-) FORNWALD J A.

(TERR-) TERRETT J A.

Irue-la-Arispe L, Hastings GA, Ruben SM, Jonak ZL, Trulli SH;

Formwald JA, Terrett JA;

WPI: 2001-025136/03.

P-PSDB: AAB50002.

METH1 and METH2 polynucleotides and encoded polypeptides, used to

inhibit angiogenesis in the treatment of disorders such as cancer,

rheumatoid arthritis and psoriasis -

Claim 7; Fig 1: 768pp; English.

The present sequence is the coding sequence of human METH1 (ME for

metalloproteinase and TR for thrombospondin). METH1 can be used for

inhibiting angiogenesis in an individual, and for treating cancer, benign

tumors, an ocular angiogenic disease, rheumatoid arthritis, psoriasis,

delayed wound healing, endometriosis, scleroderma, trachoma, vascular

adhesions, myocardial angiogenesis, coronary collateral, cerebral

collaterals, arteriovenous malformations, ischemic limb angiogenesis,

Osler-Webber syndrome, plaque neovascularisation, telangiectasia,

haemophilic joints, angiodysplasia, fibromuscular dysplasia, wound

granulation, Crohn's disease or atherosclerosis. METH1 can also be used

in birth control. METH1 can also be used in diagnostic methods for the

prognosis of cancer. The METH1 gene is located on chromosome 21q21.

Sequence 3261 BP; 844 A; 789 C; 924 G; 700 T; 4 other;

Query Match 20.6%: Score 280.2; DB 22; Length 3261;

Best Local Similarity 62.3%: Pred. No. 2.7e-59;

Matches 457; Conservative 0; Mismatches 273; Indels 3; Gaps 1;

OY 626 GACCCGGAAGAGCCGCTTTGCTTCACTAGTAGATTGTTGGAGACAGCTGGTG 685  
 DB 686 GAAGCATAGAAAGAGGATTTGTCCAGTCACCGCTATGTGGAACCATGCTTGTG 745  
 OY 686 CAGATGACAAAGATGGCCGATTCACAGGTGGGGGCTAAAGGCTACCTGCTAACAGTGA 745  
 DB 746 CAGACCAAGTCGAGGATTCACAGGAGTGTCTTAAGCATTAACCTTCACCTTGT 805





PD 17-MAY-2001.  
 XX 10-NOV-2000; 2000MO-JF07917.  
 XX 11-NOV-1999; 99JP-0321740.  
 PR 16-MAY-2000; 2000JP-0144020.  
 XX (YAMA) YAMANOUCHI PHARM CO LTD.  
 PA (KAZU-) KAZUSA DNA RES INST.  
 XX  
 PI Yamaji N, Nishimura K, Abe K, Ohara O, Nagase T, Nomura N;  
 DR WPI: 2001-343602/36.  
 DR P-PSDB: AAG62299.  
 XX  
 PT Metalloprotease with aggrecanase activity for treating joint diseases  
 XX especially osteoarthritis  
 PS Example 2; Page 60-61; 85pp; Japanese.  
 CC This invention relates to a metalloprotease with aggrecanase activity.  
 CC The invention includes protein and DNA sequences of the metalloprotease,  
 CC vectors containing the DNA, host cells transformed by the vectors, and  
 CC antibodies directed against the metalloprotease. The antibodies, protein  
 CC and DNA sequences can be used in the treatment and prevention of joint  
 CC diseases, particularly osteoarthritis. The treatment may result in  
 CC osteopathic and antiarthritic activity. The present sequence represents  
 CC cDNA encoding the metalloprotease termed MDT56.  
 XX

Sequence 2853 BP; 522 A; 920 C; 903 G; 508 T; 0 other:

Query Match 20.5%; Score 279.2; DB 22; Length 2853;  
 Best Local Similarity 63.4%; Pred. No. 4.5e-59;  
 Matches 444; Conservative 0; Mismatches 253; Indels 3; Gaps 1;

QY 618 CAGCCCCAGAGCCCGAAGCCGCTTGTCTTCACTGATGATTTGTGGAGACACT 677  
 DB 609 CCGGCGGAGGTCGTGGGCGCGCAAGCGTTTGTCTTATCCCGCGTACGTGGAACCT 668  
 QY 678 GGTGTGGCAGATGACAGATGCGCGCATTCAGGTCGGGGTAAGCCCTACTACTCT 737  
 DB 669 GGTGTGGCAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 728  
 QY 738 AACAGTGTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 797  
 DB 729 GACGCTGCTGCGCAAGCGCGCGCGCATCTACCGCATCTCCAAACCCCATCAA 788  
 QY 798 CTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 857  
 DB 789 CATGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 848  
 QY 858 CAGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 917  
 DB 849 CAATGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 908  
 QY 918 GACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 977  
 DB 909 TGACAAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 968  
 QY 978 AGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1037  
 DB 969 AGCCAAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1028  
 QY 1038 GAGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1097  
 DB 1029 AAGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1088  
 QY 1098 GGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1157  
 DB 1089 GGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1148  
 QY 1158 TTTGAGCACTCTCTGCGCATGATGATGATGATGATGATGATGATGATGATGATGAT 1217

DB 1149 GCTCCGAGCC--AACACATGATGTCCCGAGCCCTCATTCACAGATGACCCGCAACCC 1205  
 QY 1218 CTGGTCCCGCTGACAGTGGCGGCTTATCATCATGATCTCTGAGACATGGGCACTG 1277  
 DB 1206 CTGGTCCCGCTGACAGTGGCGGCTTATCATCATGATCTCTGAGACAGGGGCACTG 1265  
 QY 1278 TCTCTTACGACAAACAGAGGCTTCATTCATCTGCTGTG 1317  
 DB 1266 CCTCTGTGACCAACCCAGACAGCCCATCTCCCTGCCGAG 1305

RESULT 15  
 AAD35569  
 ID AAD35569 standard; cDNA: 2853 BP.  
 XX  
 AC AAD35569;  
 XX  
 DT 26-JUL-2002 (first entry)  
 XX  
 DE Human protease cDNA #2.  
 XX

KW Human: novel human protein: NHP; protease; biological disorder: obesity;  
 KW high blood pressure; arthritis; connective tissue disorder; infertility;  
 KW gene therapy; enzyme; gene; ss.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..2853  
 FT /\*tag= a  
 FT /product= "Human protease #2"

MO200226949-A2.  
 PD 04-APR-2002.  
 XX  
 PF 27-SEP-2001; 2001WO-US30350.  
 PR 29-SEP-2000; 2000US-23689P.  
 XX  
 PA (LEXI-) LEXICON GENETICS INC.  
 XX  
 PI Fridge CJ, Hlilun E;  
 XX  
 DR WPI: 2002-372123/40.  
 DR P-PSDB: AAD22541.  
 XX

Novel nucleic acid encoding a human protease, useful as a hybridization  
 probe for screening libraries and assessing gene expression patterns -  
 Claim 1: Page 35-36; 41pp; English.  
 XX  
 CC The present sequence is a cDNA encoding novel human protein (NHP),  
 CC human protease. NHPs share structural similarity with animal proteases  
 CC particularly zinc metalloproteases. Sequences of the invention are  
 CC useful in therapeutic, diagnostic and pharmacogenomic applications.  
 CC NHP polynucleotides are used as hybridisation probes for screening  
 CC libraries and assessing gene expression patterns. They can also be  
 CC used for treating related biological disorders such as obesity, high  
 CC blood pressure, arthritis, connective tissue disorders and infertility.  
 CC They are also used in gene therapy.  
 XX

Sequence 2853 BP; 522 A; 920 C; 903 G; 508 T; 0 other:  
 Query Match 20.5%; Score 279.2; DB 24; Length 2853;  
 Best Local Similarity 63.4%; Pred. No. 4.5e-59;  
 Matches 444; Conservative 0; Mismatches 253; Indels 3; Gaps 1;

QY 618 CAGCCCCAGAGCCCGAAGCCGCTTGTCTTCACTGATGATTTGTGGAGACACT 677  
 DB 609 CCGGCGGAGGTCGTGGGCGCGCAAGCGTTTGTCTTATCCCGCGTACGTGAGACCT 668  
 QY 678 GGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 737

Db 669 GGTGGTCGGGGAGAGTCATGCTCAAGTTCCAGCGGGAGACTGGAAACATTATCTGCT 728  
QY 738 AACAGTATGGGAGCAGACGCCAAGGCTTCAAGCACCCAGCATCCGCAATCCTGTCAAG 797  
Db 729 GAGCTGTGCTGGCAACGGCGCGGACTTACCGCATCCAGCATCTCAACCCCATCAAA 788  
QY 798 CTGGTGTGACTCGGCTAGTGTATCCTGGGGTCAAGGAGAGGGGCCCAAGTGGGGCC 857  
Db 789 CATCGTTGTGTCAAGTGTGCTGCTTCTTAGAGATCGTGAATCCGGGCCCAAGTCAACGG 848  
QY 858 CAGTGTGCCCCAGACCCCTTGCGGAGCTTGTGCTGGCAGACGGGGCTCAACACCCCTGA 917  
Db 849 CAATGGGGCCCTGACGCTGCGCAACTTCTGCTGGCAGAGAAAGCTGAAACAAAGTGAG 908  
QY 918 GGACTGGAGCCCTGACACTTTGACACAGCCATCTGTTACCCGTCAAGACCTGTGTGG 977  
Db 909 TGCACAAAGCAACCCGAGTACTGGGACACTGCGCATCTTCAACCAAGGAGGACCTGTGTGG 968  
QY 978 AGTCTCCACTTGGGACAGCCTGGGTATGGCTGATGTGGCACCGTCTGTACCCGGCTCG 1037  
Db 969 AGCCACACACTGTGACACACCCCTGGGATGGCTGATGTGGGTACCATGTGTGACCCCAAGAG 1028  
QY 1038 GAGCTGTGCATTGTGGAGATGATGGCTCCAGTCAAGCTTACTGCTGTCTATGAAT 1097  
Db 1029 AAGCTGCTCTGTGATGAGAGAGATGGCTTCCATCAGCCTTCACCACTGCCACAGAGCT 1088  
QY 1098 GGGTCATGCTTCAACATGCTCCATGACAACTCCAGCCATGATCAGTTTGAATGGGCC 1157  
Db 1089 GGGCCACAGTGTACATGCCCCATGACAAATGTGAAGTCTGTGAGAGAGTGTGGGAA 1148  
QY 1158 TTTGAGCACCTCTGCGCATGTATGGCCCCCTGTGATGGCTCATGTGATCTGAGAGGCC 1217  
Db 1149 GCTCCGAGCC---AACCACATGATGTCCCCGAGCCCTCATCCAGATCGACCGTCCAAACC 1205  
QY 1218 CTGGTCCCCCTGAGTGGCCGCTTCACTCACTGACTTCCCTGGACAATGGCTATGGGACTG 1277  
Db 1206 CTGGTCAAGCTGTGAGTGTGCTCATTCATCAACGACTTCCCTGGACAGCGGGCAGGTTGACTG 1265  
QY 1278 TCTCTAGACAACACAGAGGCTCCATTGCATCTGCCTGTG 1317  
Db 1266 CCTCCTGGACCAACCGACGAGGCCCATCTCCCTGCCCGAG 1305

Search completed: April 25, 2003, 16:50:38  
Job time : 624.069 secs

GenCore version 5.1.4-p5-4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: April 25, 2003, 13:22:53 ; Search time 7175.05 seconds

(without alignments)  
5512.258 Million cell updates/sec

Title: US-10-050-200-1

Perfect score: 1359  
Sequence: 1 gaattcgcatgtccagac.....atgacaagggtagtcgac 1359Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapept 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rdg:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1320	97.1	4301	9 AB014588	AB014588 Homo sapi
2	1318.4	97.0	4193	9 AF148213	AF148213 Homo sapi
3	1316.8	96.9	4407	6 AX376284	AX376284 Sequence
4	995.6	73.3	3321	10 BC027773	BC027773 Mus muscu
5	957.8	70.5	1991	9 BC030812	BC030812 Homo sapi
6	634.4	46.7	10766	9 AY044847	AY044847 Homo sapi
7	634.4	46.7	13512	9 AL590714	AL590714 Human DNA
8	634.4	46.7	167863	2 AL590651	AL590651 Homo sapi
9	543.4	40.0	1251	10 AB042271	AB042271 Rattus no
10	434.8	32.0	211772	10 AC084821	AC084821 Mus muscu
11	434.8	32.0	227141	2 AC093409	AC093409 Mus muscu
12	397.2	29.2	84868	2 AC099236	AC099236 Rattus no
13	288.8	21.3	576	4 AF368321	AF368321 Equus cab
14	281.2	20.7	4878	10 AF149118	AF149118 Rattus no
15	280.2	20.6	2853	6 E58656	E58656 Novel metal
16	280.2	20.6	3430	9 AF060152	AF060152 Homo sapi
17	280.2	20.6	4014	6 E29406	E29406 Novel Integ
18	280.2	20.6	4309	9 AB037767	AB037767 Homo sapi
19	280.2	20.6	4659	9 AF207664	AF207664 Homo sapi
20	280.2	20.6	4670	9 BC036515	BC036515 Homo sapi
21	279.2	20.5	2853	9 HSA315733	HSA315733 Homo sapi
22	279.2	20.5	2930	6 AX342635	AX342635 Sequence
23	279.2	20.5	2937	6 AX319860	AX319860 Sequence
24	278.6	20.5	4447	9 AF170084	AF170084 Homo sapi
25	277.4	20.4	2184	6 E29966	E29966 Human ADAMT
26	276.4	20.3	2857	10 AF304446	AF304446 Rattus no
27	276.4	20.3	2857	10 D67076	D67076 Mouse mRNA
28	271.2	20.0	600	4 AF192770	AF192770 Bos tauru
29	249.8	18.4	3711	9 AF060153	AF060153 Homo sapi
30	248.2	18.3	2670	6 E55282	E55282 Novel metal
31	248.2	18.3	2670	6 E58655	E58655 Novel metal
32	246.4	18.1	3638	10 AF175282	AF175282 Mus muscu
33	246.4	16.8	739	9 AF175283	AF175283 Mus muscu
34	180.6	13.3	3002	10 AF140673	AF140673 Mus muscu
35	178.6	13.1	3353	10 AF449483	AF449483 Mus muscu
36	176	13.0	1365	10 AB042272	AB042272 Rattus no
37	175.8	12.9	2554	9 AK094057	AK094057 Homo sapi
38	175.8	12.9	5533	9 AF142099	AF142099 Homo sapi
39	143.6	10.6	118241	9 AP001599	AP001599 Homo sapi
40	143.6	10.6	171038	9 AP001598	AP001598 Homo sapi
41	143.6	10.6	340000	9 AP001697	AP001697 Homo sapi
42	139.8	10.3	182656	2 AC101990	AC101990 Mus muscu
43	137.8	10.1	9248	10 AB001735	AB001735 Mus muscu
44	137.4	10.1	3179	6 AX481383	AX481383 Sequence
45	137.4	10.1	3179	9 AF140675	AF140675 Homo sapi

## ALIGNMENTS

RESULT 1  
AB014588  
LOCUS AB014588 4301 bp mRNA  
DEFINITION Homo sapiens mRNA for KIAA0688 protein, complete cds.  
ACCESSION AB014588  
VERSION AB014588.1 GI:3327189  
KEYWORDS  
SOURCE Homo sapiens adult male brain cDNA to mRNA, clone: jlb:pluescript11  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (sites) Ishikawa, K., Nagase, T., Suyama, M., Miyajima, N., Tanaka, A.,  
Kotani, H., Nomura, N. and Ohara, O.

TITLE Prediction of the coding sequences of unidentified human genes. X.  
The complete sequences of 100 new cDNA clones from brain which can  
code for large proteins in vitro

JOURNAL DNA Res. 5 (3), 169-176 (1998)

MEDLINE 98403880

REFERENCE 2 (bases 1 to 4301)

AUTHORS Ohara, O., Suyama, M., Nagase, T. and Ishikawa, K.

TITLE Direct Submission

JOURNAL Submitted (26-MAY-1998) Osamu Ohara, Kazusa DNA Research Institute,  
Laboratory of DNA Technology; Yama 1532-3, Kisarazu, Chiba  
292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp, Tel: +81-438-52-3913,  
Fax: +81-438-52-3914)

FEATURES

source Location/Qualifiers

1..4301

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="HK03410"

/sex="male"

/tissue\_type="brain"

/clone\_lib="pBluescriptII SK plus"

/dev\_stage="adult"

1..4301

/gene="KIAA0688"

401..2914

/gene="KIAA0688"

/codon\_start=1

/product="KIAA0688 protein"

/protein\_id="BA031653.1"

/db\_xref="GI:3327190"

/translation="MSQTSHPGRGLAQMILMGAOPCLLPVLSMLVLLILLAS  
LIPSAKLPLPREIVPEKLVNGSVLPSTPARLRLQARFETLLLEDDSGV  
OVELTVQYLGAPPELLGAEPGLYLTGTINDPESVSLHMDGALLGLVQRAEL  
HLQPEGGTSPNSAGCPGAILLRKSPAGCGCMCAKPLSPSPRRARAFSLSR  
FVETLVADDDKMAAFHAGILKRYLLTVAAAKAKHPSIRNPVSLVTVRIISGE  
EGPOYPSGAOTLRSEFCAMORGILNPEISDPDHPETALLFTRODLGVSTDDTGMAD  
VGTVCDPARSAIVEDDLOSFTAAHLEHLYFNNLHNSKRCISLNPPLSTRVMA  
PVMAHVDPEEPSPCSARFITYDLNNGHGLDPEAPLHLPTFPKDDTDADQOC  
LTFPGDSRHCPOLPPCAALWCSGLNGLHACOTHSWADGTPGAPADMGHGLH  
MDLODFNIPOAGWGPMPGDCSRTCGGVGESSRCTRVPFRNGKRYCEGRTRF  
RSCNDEPTGSLTFREEOCAAYNHRDLEKSPGMDWPRYGVAPQDCKLTGC  
ARALGYVYVLEPRVVDGTPCSPDSSVGVGRCIHAGCDRIISKKKFKDKCMVCGDG  
SGCSKSSSFKFRGVNNVYITPAGATHLLVROGNGHSHIYALILPDSYALNG  
ETTLMPSTDYVLRGAVSLRTSGATASETLSGHCPLQPLTLQVLVAGNPDTRLRT  
SFVPRPSPRPDPMDLHRAQILEILRRRPWAGRK"

BASE COUNT 844 a 1267 c 1232 g 958 t

ORIGIN

Query Match 97.1% Score 1320; DB 9; Length 4301;  
Best Local Similarity 98.5% Pred. No. 1.1e-277;  
Matches 1332; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Db 698 TCCGCTGTGACAGTCGAGGGGCTGACAGTGCAGTACCTGCGGCGCCCTGAGCTGCTG 757

Qy 367 GGTGAGACAGACCTTGACACCTACCTGACTGGACCATCAATGAGATCCGAGTCGTG 426

Db 758 GGTGAGACAGACCTTGACACCTACCTGACTGGACCATCAATGAGATCCGAGTCGTG 817

Qy 427 GCATCTGTGACAGTCGAGGGGAGCCCTGTTAGCGGTTCACATATTCGGGGCTGAA 486

Db 818 GCATCTGTGACAGTCGAGGGGAGCCCTGTTAGCGGTTCACATATTCGGGGCTGAA 877

Qy 487 CTCACCTTCACCCCTGAGGAGGAGGACCCCTAACCTGCTGGGGGACCTGGGGCTAC 546

Db 878 CTCACCTTCACCCCTGAGGAGGAGGACCCCTAACCTGCTGGGGGACCTGGGGCTAC 937

Qy 547 ATCTTACGCGGAGAGAGTCTCTCCAGCGGTTCAGAGTCCCATGTGCAACCTCAAGCTCT 606

Db 938 ATCTTACGCGGAGAGAGTCTCTCCAGCGGTTCAGAGTCCCATGTGCAACCTCAAGCTCT 997

Qy 607 CTTGGAAGCCCGACAGCCCGACAGCCCGAAGAGCCAGCGCTTTGCTTCACTGATGATTT 666

Db 998 CTTGGAAGCCCGACAGCCCGACAGCCCGAAGAGCCAGCGCTTTGCTTCACTGATGATTT 1057

Qy 667 GTGAGACACCTGTGTGTGACATGACAAAGATGGCCGATTCACGCTGGGGGGCTAAAG 726

Db 1058 GTGAGACACCTGTGTGTGACATGACAAAGATGGCCGATTCACGCTGGGGGGCTAAAG 1117

Qy 727 CCTACCTGCTAACAGTGTGTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 786

Db 1118 CCTACCTGCTAACAGTGTGTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1177

Qy 787 AATCTGTACGCTGTGTGTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 846

Db 1178 AATCTGTACGCTGTGTGTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1237

Qy 847 CAAGTGGGGGCCAGTGTGTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 906

Db 1238 CAAGTGGGGGCCAGTGTGTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1297

Qy 907 AACACCCCTGAGAGTGTGTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 966

Db 1298 AACACCCCTGAGAGTGTGTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1357

Qy 967 GACCTGTGTGAGTGTGTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1026

Db 1358 GACCTGTGTGAGTGTGTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1417

Qy 1027 GACCGGGCTCGAGTGTGTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1086

Db 1418 GACCGGGCTCGAGTGTGTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1477

Qy 1087 GCTCATGAACGTGTGTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1146

Db 1478 GCTCATGAACGTGTGTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1537

Qy 1147 TTGAATGGGCTTTTGAAGCAGCTCTGCGCATGTATGAGCCCTGTATGCTATGTGAT 1206

Db 1538 TTGAATGGGCTTTTGAAGCAGCTCTGCGCATGTATGAGCCCTGTATGCTATGTGAT 1597

Qy 1207 CCTGAGAGCCCTGAGTGTGTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1266

Db 1598 CCTGAGAGCCCTGAGTGTGTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1657

Qy 1267 TATGGGCACTGTCTTTAGACAACAGAGGCTTCATTCATCTGCTGCTGAGCTGGGAGC 1326

Db 1658 TATGGGCACTGTCTTTAGACAACAGAGGCTTCATTCATCTGCTGCTGAGCTGGGAGC 1717

Qy 1327 TACAAGAGCAGAGTGTGACAAGGGGTAGTGA 1358

Db 1718 GGCMAAGACTATGATGCTGACCGCAGTGCCA 1749

RESULT 2

AF148213



LOCUS	AF148213	4193 bp	mRNA	linear	PRI 30-JUN-1999
DEFINITION	Homo sapiens aggrecanase-1 mRNA, complete cds.				
ACCESSION	AF148213				
VERSION	AF148213.1	GI:5281380			
KEYWORDS					
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens.				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Tortorella,M.D., Burn,T.C., Pratta,M.A., Abbaszade,I., Hollis,J.M., 1 (bases 1 to 4193)				
TITLE	Tortorella,M.D., Burn,T.C., Pratta,M.A., Abbaszade,I., Hollis,J.M., Rockwell,A., Yang,F., Duke,J.L., Solomon,K., George,H., Bruckner,R., Nagase,H., Itoh,Y., Ellis,D.M., Ross,O.H., Murphy,K., Hillman,M.C., Jr., Hollis,G.F., Newton,R.C., Magolda,R.L., Tzaskos,J.M. and Arner,E.C. Purification and cloning of aggrecanase-1: a member of the ADAMTS family of proteins				
JOURNAL	Science 284 (5420), 1664-1666 (1999)				
MEDLINE	99286503				
PUBMED	10356395				
REFERENCE	2 (bases 1 to 4193)				
AUTHORS	Tortorella,M.D., Burn,T.C., Pratta,M.A., Abbaszade,I., Hollis,J.M., Lu,R.-O., Rosenfeld,S.A., Copeland,R.A., Decicco,C.P., Wynn,R., Rockwell,A., Yang,F., Duke,J.L., Solomon,K., George,H., Bruckner,R., Nagase,H., Itoh,Y., Ellis,D.M., Ross,O.H., Wistwall,B.H., Murphy,K., Hillman,M.C., Jr., Hollis,G.F., Newton,R.C., Magolda,R.L., Tzaskos,J.M. and Arner,E.C. Direct Submission				
TITLE	Submitted (03-MAY-1999) Applied Biotechnology, Dupont Pharmaceuticals Company, Experimental Station, E336/237B, Wilmington, DE 19880, USA				
JOURNAL	Location/Qualifiers				
FEATURES	1..4193				
source	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
CDS	407..2920				
	/note="aggrecan degrading metalloproteinase-1; disintegrin metalloproteinase; ADMP-1; ADAMTS-4"				
	/codon_start=1				
	/product="aggrecanase-1"				
	/protein_id="AADM1499.1"				
	/db_xref="GI:5281381				
	/translation="MSQTSHPRGRLGRLWLMGAPCLLPVLSMLWLLLLLASLPARLASLPREEEIVPEKLSNGSVLPGAGARLCRLQAGETLLEEDSGGVEELAVOYLGAPELIGGAPGTYLTINGDEPSVALHMDGALLIGVYGAELHLPLEGCTPMSAGCGPANHILRRKSPASGCGPMKNVAPLSPRRARRPASTLSFVETLVYADDKMAAFHCGAKRLVLTIVAAAKAFKHPSTIRNPVLVTVLVLGGSEFGLVPPSAAGTTLSPFCAMWGRLTPESDHDFTDLILFTRDLCGVSTDTLGMALVGTIVCDPARSAIVEDGLOSAPFAAHELGHVFMLLHDSKPLSTSHVMVPMVAHVDEEPSEWCSARFIDFDNGHCLLDKPEAPLHLPTVFPKVDVADROCCLTPEPDSRHPOLPPCCALALMSGHLNHAMCQVSHMAGTGGCGAPQACMGGRCLHMDQLEDNIRPOAGSGWPMGPMGDCSRICCGGVQSSRCDTPPYRRPQCYEGRTTRRSCTGTEDEPTGSALETFREBOCAVNHNRDKLSPGPDWVPRTGVAPODCKLTCARAGVYVLEPRVVDGTPCSPDSSSVQVCRCHAGDRIIGKKFKDKMVCVGDDSGSGCGSGSEFRKPRVGYNNVYTPAGATVHILVRQDNGHRSITLALPLPGSYALNMEYTLMPSTDVLLPGAVSLRSGATASELTSGGPIAOLPTLVLVAGNQDTRLRLRFSFVRRPSPRPPTPDMDLHRRAOILELHRRRWAGK"				
BASE COUNT	807 a 1249 c 1220 g 917 t				
ORIGIN					
Query Match	97.0%; Score 1318.4; DB 9; Length 4193;				
Best Local Similarity	98.4%; Pred. No. 2.4e-277;				
Matches 1331; Conservative	0; Mismatches 21; Indels 0; Gaps 0;				
0y	7 GCGATGTCCGACAGAGGCTCGATCCCGGAGGCGCTTGGACAGGGCGCTGCTGTGGCGA 66				
Db	404 GCCATGTCCCAACAGCGCTCGCATCCCGGAGCGGCTTGGACAGGGCGCTGCTGTGGCGA 463				
0y	67 GCCCAACCTGCTCTCTGCTCCCAATTGGCGCGCTCTGCTGCTGTGGCTGTGGCTTGG 126				
Db	464 GCCCAACCTGCTCTCTGCTCCCAATTGTGGCGGCTCTCTGCTGTGGCTGTGGCTTGG 523				

OY	127	CTACTGCTGGCCCTCTCTCTCTGTCCTCTAGCCCGAGCT966CAGACGCCCTCTCCCGGGAGAG	186
Db	524	CTACTGCTGGCCCTCTCTCTCTGTCCTCTAGCCCGAGCT966CAGACGCCCTCTCCCGGGAGAG	583
OY	187	GAGATCGATTTCAGAGAGGCTCAACGGGAGGCTCTGCTGGCTCGGGACCCCTCTCC	246
Db	584	GAGATCTGTTTCCAGAGAAAGCTCAACGGGAGGCTCTGCTGGCTCGGGACCCCTCTCC	643
OY	247	AGGCTGTTTGGCCCTCTGCAAGGCTTTTGGGGAGACGCTCTACTAGAGCTGGAGCAGAC	306
Db	644	AGGCTGTTTGGCCCTCTGCAAGGCTTTTGGGGAGAGGCTCTACTAGAGCTGGAGCAGAC	703
OY	307	TCCGGTGTGACAGTCCAGGGGGCTTGACAGTGCAGTACCTTGGCCAGCGCTGAGCTGCTG	366
Db	704	TCCGGTGTGACAGGCTCGAGGGGGCTTGACAGTGCAGTACCTTGGCCAGCGCTGAGCTGCTG	763
OY	367	GGTGAGACAGAGCCTGGCACCCTACCTGAGCTGGCAGCATCAATGAGAGATCCGGAGTCCGTG	426
Db	764	GGTGAGACAGAGCCTGGCACCCTACCTGAGCTGGCAGCATCAATGAGAGATCCGGAGTCCGTG	823
OY	427	GCAATCTGTCAGTGGGATTTGGGGGAGGCCCTTTTGGCGTGTACATATTCGGGGGCTGAA	486
Db	824	GCAATCTGTCAGTGGGATTTGGGGGAGGCCCTTTTGGCGTGTACATATATCGGGGGGCTGAA	883
OY	487	CTCCACCTCCAGGCCCTTGAGAGGAGACCCCTAAGCTGCTGGGGGAGCTTG6GGGCTAC	546
Db	884	CTCCACCTCCAGCCCCCTGAGAGGAGAGCACCCCTAAGCTGCTGGGGGAGCTTG6GGGCTAC	943
OY	547	ATTCACGCGCGGAAGATTCCTGGCAGCGGTCAGAGTCCCATGTCGACAGCTCAAGGCTCT	606
Db	944	ATTCCTACGCGCGGAAGATTCCTGGCAGCGGTCAGAGTCCCATGTCGACAGCTCAAGGCTCT	1003
OY	607	CTTGAAGGCCCCACACCCCGACAGACCCCGAGAGGCCAACCGCTTGTGCTTACAGTAGTAGATT	666
Db	1004	CTTGAAGGCCCCACACCCCGACAGACCCCGAGAGGCCAACCGCTTGTGCTTACAGTAGTAGATT	1063
OY	667	GTCGAGACACTGTGTGTGGCAGATGACAMAGTGGCCGATTCACAGGTGCGGGGCTAAAG	726
Db	1064	GTCGAGACACTGTGTGTGGCAGATGACAMAGTGGCCGATTCACAGGTGCGGGGCTAAAG	1123
OY	727	CGATACCGCTCAACAGTAGTAGTGGCAGGACGCCAACGGCTTCACAGCACCCAGCAATCCGC	786
Db	1124	CGATACCGCTCAACAGTAGTAGTGGCAGGACGCCAACGGCTTCACAGCACCCAGCAATCCGC	1183
OY	787	AATCTGTGACAGCTTGAGTGTGAGTCTGGGCTAGTATCTGGGGTCAAGGCGAGGAGGGGCC	846
Db	1184	AATCTGTGACAGCTTGAGTGTGAGTCTGGGCTAGTATCTGGGGTCAAGGCGAGGAGGGGCC	1243
OY	847	CAAGTGGGGGCCAGTGTCTGCCAGACACCCTGCGCAGCTTGTGTGCTGGCAGCGGGGCTTC	906
Db	1244	CAAGTGGGGGCCAGTGTCTGCCAGACACCCTGCGCAGCTTGTGTGCTGGCAGCGGGGCTTC	1303
OY	907	AACACCCCTGAGGACCTGGACCCCTGACACNATTTGACACAGCCATTCCTGTTACCCGCTAG	966
Db	1304	AACACCCCTGAGGACCTGGACCCCTGACACNATTTGACACAGCCATTCCTGTTACCCGCTAG	1363
OY	967	GACCTGTGTGAGTCTCCACTTGGCAGACGCTGGATGATGGCTGATGTGGGACCGCTGT	1028
Db	1364	GACCTGTGTGAGTCTCCACTTGGCAGACGCTGGATGATGGCTGATGTGGGACCGCTGT	1422
OY	1027	GACCCGGCTCGGAGCTGTGCCATTGTGAGAGATGATGGCTCCAGTCAAGCTTCACTGCT	1086
Db	1424	GACCCGGCTCGGAGCTGTGCCATTGTGAGAGATGATGGCTCCAGTCAAGCTTCACTGCT	1483
OY	1087	GCTATGAACTGGGTCAATGCTTCAACATGTCCTCAATGACACATCCCAAGCATGATCACT	1146
Db	1484	GCTATGAACTGGGTCAATGCTTCAACATGTCCTCAATGACACATCCCAAGCATGATCACT	1543
OY	1147	TTGATATGGGCTTTTGACACACTCTCGGCATATGATGGCCCTGTGATGCTATGTGAT	1206
Db	1544	TTGATATGGGCTTTTGACACACTCTCGGCATATGATGGCCCTGTGATGCTATGTGAT	1603
OY	1207	CCTGAGAGGCCCTGTGCTCCCTGCGAGTCCCGCTTCACTGACTTACTGTGACATGGC	1266

```
|||||
Db 1604 CCTGAGAGACCCCTGGTCCCGCCCTGACAGTGGCCGCTTATCATCTGACTGCTGGACAATGGC 1663
QY 1267 TTGTGGGACATGCTCTTTAGACAAACAGAGGCTCATTCGATCTGCTGACTGGGAC 1326
Db 1664 TATGGGCACTGCTCTTGAACAACAGAGGCTCATTCGATCTGCTGACTTTCCT 1723
QY 1327 TACAAGAGACAGATACAAAGGGTACGTGA 1358
Db 1724 GGCAAGGACTATGATGCTGACCGCCAGTCCCA 1755

RESULT 3
AX376284 4407 bp DNA Linear PAT 01-MAR-2002
LOCUS Sequence 351 from Patent WO0168848.
DEFINITION AX376284
ACCESSION AX376284.1 GI:19170533
VERSION
KEYWORDS
SOURCE
ORGANISM human.
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 Baker, K.P., Chen, J., Desnoyers, L., Goddard, A., Godowski, P.J.,
Gurney, A.L., Pan, J., Smith, V., Watanabe, C.K., Wood, W.I. and
Zhang, Z.
Secreted and transmembrane polypeptides and nucleic acids encoding
the same
JOURNAL Patent: WO 0168848-A 351 20-SEP-2001;
Genentech, Inc. (US)
FEATURES
source 1. 4407
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 932 a 1275 c 1241 g 959 t
ORIGIN
Query Match 96.9%; Score 1316.8; DB 6; Length 4407;
Best Local Similarity 98.4%; Pred. No. 5.4e-277;
Matches 1330; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
```

```
QY 487 CTCACCTCCACCCCTGGAGAGGACACCCCTAACTCTGTGGGGACTGGGCTCAG 546
Db 896 CTCACCTCCACCCCTGGAGAGGACACCCCTAACTCTGTGGGGAGNCTGGGGCTCAG 955
QY 547 ATCTAACGCCGGAAGAGTCTCCAGCGGTCAAGTCCCATGTGCAGCTCAAGCTCTC 606
Db 956 ATCTAACGCCGGAAGAGTCTCCAGCGGTCAAGTCCCATGTGCAGCTCAAGCTCTC 1015
QY 607 CTGGAGGCCGCCAGCCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 666
Db 1016 CTGGAGGCCGCCAGCCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1075
QY 667 GTGAGACACTGGTGGTGGAGATGACAAGATGGCCGATTCACAGTCCGGGCTAAAG 726
Db 1076 GTGAGACACTGGTGGTGGAGATGACAAGATGGCCGATTCACAGTCCGGGCTAAAG 1135
QY 727 CGCTACTGCTAACAGTATGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 786
Db 1136 CGCTACTGCTAACAGTATGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1195
QY 787 ATCTCTGACCTTGTGTGTGATCGGCTAGTATCTGGGGTACAGGAGAGGAGGAGG 846
Db 1196 ATCTCTGACCTTGTGTGTGATCGGCTAGTATCTGGGGTACAGGAGAGGAGGAGG 1255
QY 847 CAAGTGGGGCCGAGTGTCCAGACCCCTGCGACCTTGTGCTGGCAGCGGGGCTTC 906
Db 1256 CAAGTGGGGCCGAGTGTCCAGACCCCTGCGACCTTGTGCTGGCAGCGGGGCTTC 1315
QY 907 AACACCCCTGAGAGTCTCGACCCCTGACCACTTTGACACAGCCATTCTGTTTACCCTGAG 966
Db 1316 AACACCCCTGAGAGTCTCGACCCCTGACCACTTTGACACAGCCATTCTGTTTACCCTGAG 1375
QY 967 GACCTGTGGAGTGTCCACTGGAGACAGCGTGGTATGGATGGTGGGACCGCTGT 1026
Db 1376 GACCTGTGGAGTGTCCACTGGAGACAGCGTGGTATGGATGGTGGGACCGCTGT 1435
QY 1027 GACCGGCTGGAGTGTCCACTGGAGAGATGGATGGGCTTCAGTCAAGCTTCACTGCT 1086
Db 1436 GACCGGCTGGAGTGTCCACTGGAGAGATGGATGGGCTTCAGTCAAGCTTCACTGCT 1495
QY 1087 GCTCATGAACCTGGTATCTTCAACATGCTTCATGAACCTGAACCTGAACCTGAACCT 1146
Db 1496 GCTCATGAACCTGGTATCTTCAACATGCTTCATGAACCTGAACCTGAACCTGAACCT 1555
QY 1147 TTGAATGGGCTTTGAGACACTCTGCGCATGTGATGGCCCTGTGATGGCTATGGAT 1206
Db 1556 TTGAATGGGCTTTGAGACACTCTGCGCATGTGATGGCCCTGTGATGGCTATGGAT 1615
QY 1207 CCTGAGAGACCCCTGGTCCCTGTGACAGTGGCCGCTTATCACTGACTTCTGGACAATGGC 1266
Db 1616 CCTGAGAGACCCCTGGTCCCTGTGACAGTGGCCGCTTATCACTGACTTCTGGACAATGGC 1675
QY 1267 TATGGGCACTGTCTTGAACAACAGAGAGGCTCATTCGATTCGCTGGTGGGAGC 1326
Db 1676 TATGGGCACTGTCTTGAACAACAGAGGCTCATTCGATTCGCTGGTGGGAGC 1735
QY 1327 TACAAGAGACAGATGACAAGGGGTAGGTGA 1358
Db 1736 GGCAAGGACTATGATGCTGACCGCCAGTCCCA 1767

RESULT 4
BC027773 3321 bp mRNA Linear ROD 07-AUG-2002
LOCUS BC027773
DEFINITION Mus musculus. Similar to a disintegrin-like and metalloprotease
(reprolysin type) with chondrospindin type 1 motif, 4, clone
MGC:38401 IMAGE:5345809, mRNA, complete cds.
ACCESSION BC027773
VERSION BC027773.1 GI:20381316
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
```





QY 907 AACACCCCTGAGAGCTCGACCTGACACTTGTGACACAGCCATCTGTTTACCCGTCAG 966  
|||||  
DB 1324 AACACCCCTGAGAGCTCGACCTGACACTTGTGACACAGCCATCTGTTTACCCGTCAG 1383  
QY 967 G 967  
DB 1384 G 1384

RESULT 6  
AY044847 10766 bp DNA linear PRI 18-SEP-2001  
LOCUS Homo sapiens aggreccanase 1 (ADAMTS4) gene, complete cds.  
DEFINITION AY044847  
ACCESSION AY044847.1 GI:15667234  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 10766)  
Sawaji.Y., Nagase.H., Saklatvala.J. and Clark.A.R.  
ADAMTS-4 genomic locus  
JOURNAL Unpublished  
2 (bases 1 to 10766)  
Sawaji.Y., Nagase.H., Saklatvala.J. and Clark.A.R.  
Direct Submission  
TITLE Submitted (10-JUL-2001) Kennedy Institute of Rheumatology Division,  
Imperial College, Aspenlea Road, London W6 8LH, UK  
JOURNAL Location/Qualifiers  
1. 10766  
FEATURES  
source  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/map="1q21-q23"  
/feature="10350"  
/gene="ADAMTS4"  
join(1181..2219,3334..3657,3911..4043,4579..4749,  
5994..6280,6389..6575,6827..7002,7975..8150,8651..>10350)  
/gene="ADAMTS4"  
/product="aggreccanase 1"  
join(1587..2219,3334..3657,3911..4043,4579..4749,  
5994..6280,6389..6575,6827..7002,7975..8150,8651..9077)  
/gene="ADAMTS4"  
/EC\_number="3.4.24"  
/note="metalloproteinase: ADAMTS-4; ADMP-1;  
aggreccan-degrading enzyme"  
/codon\_start=1  
/product="aggreccanase 1"  
/protein\_id="AA102262.1"  
/db\_xref="GI:15667235"  
/translation="MSOTGSHPRGLAGRWLMGAOPCLIPVPIPLVWMLILLLAS  
LLPSARLASPLPREELVPEPKLNSVLPSSAPRRLICRIQAEFTLLLEEDQSGV  
OYEGLVTOIGAPELGAEFGTYLTGTINDPESVSLHMDGALLGLVLYRAEL  
HLQPLEGTPNSAGPGVAILLRKSPASOGPCWVKAPLSPSPRPARKFVLSIR  
FETLVVADDDKMAAFHAGLKRRLYLTVMAAAKAKHPSIRNPVSLVLYRLILSSE  
ESPOVPSAAOTLRFSCAMORGLNTPEDSDPHFTALIFETODLGGVSTCITLMAD  
VGTCDPARSCAIVEDDGLQSAFTAAHGLHYFNMLHNSKPCISLNGPLSHRYMA  
PYMAHYDEBEPMSPSARITPDLNLTGCHCLLDKRPBALHLPVFPKGDYADNOCQ  
LTFGDSRRCQPLPPCAALNSGHLNGHACOTKHSVPADLTPCGPADACGGRCLH  
MDQLDFNIPQAGWGPMGDCSRTSGGGVQFSSRCDTRVPSPNGGAYCEGRTFR  
RSCNTEDCPTGSLTFREBOCAAYNHRTDLFKSPGPMVRYTGVAPODCKLTQC  
AOLATYVLEPRVVDGPGSPSSSVCYCGCHAGCDRIISKKKPKCWGGDR  
SGCSKOSGFRKRYGNVYTPAGATHTIIVROGNGPHRSITLALIKLPSGVALNG  
ETTLMPSPDVLVPGAVSLRYSGATTAASSTLSGGRPLAOPLLVQVAGNPDTLRY  
SFTVPRPPTSTPPTPQDMLHRAQILLETLRKRPWAGR"

misc\_difference  
6968  
/gene="ADAMTS4"  
/note="compared to GenBank Accession Number AF148213"  
/replace="g"  
misc\_difference  
8107  
/gene="ADAMTS4"  
/note="compared to GenBank Accession Number AF148213"

/replace="g"  
misc\_difference  
9544-9545  
/gene="ADAMTS4"  
/note="compared to GenBank Accession Number AF148213"  
/replace="lg"  
misc\_difference  
9665  
/gene="ADAMTS4"  
/note="compared to GenBank Accession Number AF148213"  
/replace="c"  
misc\_difference  
9668..9669  
/gene="ADAMTS4"  
/note="compared to GenBank Accession Number AF148213"  
/replace=""  
misc\_difference  
9694  
/gene="ADAMTS4"  
/note="compared to GenBank Accession Number AF148213"  
/replace=""  
BASE COUNT 2326 a 2956 c 2894 g 2590 t  
ORIGIN  
Query Match 46.7% Score 634.4 DB 9 Length 10766:  
Best Local Similarity 99.8% Pred. No. 3.8e-128;  
Matches 635; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 7 GCCATGTCCACAGACGCTCGCATCCCGGAGGGCTTGGACGGCGCTGTGGGA 66  
DB 1584 GCCATGTCCACAGACGCTCGCATCCCGGAGGGCTTGGACGGCGCTGTGGGA 1643  
QY 67 GCCCAACCTGCTCCTGCTCCCATTTGCGGCTCTCTGCTGTGCTGCTTCTG 126  
DB 1644 GCCCAACCTGCTCCTGCTCCCATTTGCGGCTCTCTGCTGTGCTGCTTCTG 1703  
QY 127 CTACGCTGCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 186  
DB 1704 CTACGCTGCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1763  
QY 187 GAGATGCTGTTTCAGAGAGGCTCAACGCGACGCTGCTGCTGCTGCTGCTGCTG 246  
DB 1764 GAGATGCTGTTTCAGAGAGGCTCAACGCGACGCTGCTGCTGCTGCTGCTGCTG 1823  
QY 247 AGGCTGTTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 306  
DB 1824 AGGCTGTTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1883  
QY 307 TCCGGGTGTCAGAGGCTGACAGCTGACAGTACCTGCTGCTGCTGCTGCTGCTG 366  
DB 1884 TCCGGGTGTCAGAGGCTGACAGCTGACAGTACCTGCTGCTGCTGCTGCTGCTG 1943  
QY 367 GGTGAGCAGAGGCTGACCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 426  
DB 1944 GGTGAGCAGAGGCTGACCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2003  
QY 427 GCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 486  
DB 2004 GCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2063  
QY 487 CTCACCTCAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 546  
DB 2064 CTCACCTCAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2123  
QY 547 ATCTCTAGGCGGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 606  
DB 2124 ATCTCTAGGCGGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2183  
QY 607 CTGGAAGCCCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 642  
DB 2184 CTGGAAGCCCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2219

RESULT 7  
AL590714/c 133512 bp DNA linear PRI 05-APR-2002  
LOCUS AL590714/c  
DEFINITION Human DNA sequence from clone RP11-297K8 on chromosome 1, complete

sequence.  
 AL590714 GI:20068455  
 VERSION AL590714.27  
 HTG.  
 KEYWORDS human.  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1  
 Lad, H.  
 Direct Submission  
 Submitted (04-APR-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
 On Apr 7, 2002 this sequence version replaced gi:19572754.  
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
 This sequence was finished as follows otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., paired quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: EM: EMBL; SW: SWISSPROT; TR: TrEMBL; WP: WORMPEP; information on the WORMPEP database can be found at  
[http://www.sanger.ac.uk/projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping group. Further information can be found at  
<http://www.sanger.ac.uk/HGP/Chr1>  
 RP11-297K8 is from the library RP11-11.2 constructed by the group of Pieper de Jong. For further details see  
<http://www.chori.org/bacpac/home.htm>  
 VECTOR: pBACe3.6.

FEATURES  
 source location/Qualifiers  
 1..133512  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="1"  
 /clone="RP11-297K8"  
 /clone.lib="RP11-11.2"  
 BASE COUNT 33552 a 31947 c 32880 g 35133 t  
 ORIGIN

Query Match 46.7%; Score 634.4; DB 9; Length 133512;  
 Best local similarity 99.8%; Pred. No. 3e-128;  
 Matches 635; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 GCCATATCCACAGACGGCTCCGATCCCGGAGGGCTTGCGAGGGCTGTGGGA 66  
 |||||||  
 DB 53810 GCCATATCCACAGACGGCTCCGATCCCGGAGGGCTTGCGAGGGCTGTGGGA 53751  
 |||||||

QY 67 GCCCAACCGTGGCTGCTCCGATGGCGGCTCTGCTGGTGGTGGCTTGG 126  
 |||||||  
 DB 53750 GCCCAACCGTGGCTGCTCCGATGGCGGCTCTGCTGGTGGTGGCTTGG 53691  
 |||||||

QY 127 CTACTGCTGGCTCTCTGCTGAGCCGCGGTCAGGCGCCCTCCCGGAGAG 186  
 |||||||  
 DB 53690 CTACTGCTGGCTCTCTGCTGAGCCGCGGTCAGGCGCCCTCCCGGAGAG 53631  
 |||||||

QY 187 GAGATCGTTTTCAGAGAGCTCAAGCGAGCGTCTGCTGGCTGGGACCCCTGCC 246  
 |||||||  
 DB 53630 GAGATCGTTTTCAGAGAGCTCAAGCGAGCGTCTGCTGGCTGGGACCCCTGCC 53571  
 |||||||

QY 247 AGGCTGTGCTGGAGGCTTGGGAGAGCGCTCTACTAGAGCTGAGAGCGAGC 306  
 |||||||

DB 53570 AGGCTGTGCTGGAGGCTTTGGGAGAGCGCTCTACTAGAGCTGAGCAGAC 53511  
 |||||||  
 QY 307 TCCGCTGTGACAGTGCAGAGGGCTGACAGTACCTGGGCGAGCGCTGAGCTGTG 366  
 |||||||  
 DB 53510 TCCGCTGTGACAGTGCAGAGGGCTGACAGTACCTGGGCGAGCGCTGAGCTGTG 53451  
 |||||||

QY 367 GGTGAGCAGAGCGTGGACCTTACTGCTGACCATCATATGAGAGATCCGAGTGGTG 426  
 |||||||  
 DB 53450 GGTGAGCAGAGCGTGGACCTTACTGCTGACCATCATATGAGAGATCCGAGTGGTG 53391  
 |||||||

QY 427 GCATCTCTGACCTGGAGTGGGAGCGGCTTGAAGCTTCAATATGAGGCGCTGAA 486  
 |||||||  
 DB 53390 GCATCTCTGACCTGGAGTGGGAGCGGCTTGAAGCTTCAATATGAGGCGCTGAA 53331  
 |||||||

QY 487 CTCACCTTCACGCGCTGAGGAGGAGGACCCCTTAACCTGTGGGGAGCTGGCTAC 546  
 |||||||  
 DB 53330 CTCACCTTCACGCGCTGAGGAGGAGGACCCCTTAACCTGTGGGGAGCTGGCTAC 53271  
 |||||||

QY 547 ATCTTACGCGCGGAGAGTCTGCGACGGTCAAGTCCATGTGCAAGCTCAAGCTCTCT 606  
 |||||||  
 DB 53270 ATCTTACGCGCGGAGAGTCTGCGACGGTCAAGTCCATGTGCAAGCTCAAGCTCTCT 53211  
 |||||||

QY 607 CTTGGAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAG 642  
 |||||||  
 DB 53210 CTTGGAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAG 53175  
 |||||||

RESULT 8  
 AL590651 167863 bp DNA linear HTG 18-AUG-2001  
 LOCUS  
 DEFINITION Homo sapiens chromosome 1 clone RP11-137A12. \*\*\* SEQUENCING IN PROGRESS \*\*\*. 9 unordered pieces.  
 ACCESSION AL590651  
 VERSION AL590651.4 GI:13990236  
 HTG: HTGS\_PHASE1; HTGS\_CANCELLED.  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 Harrison, E.  
 1 (bases 1 to 167863)  
 JOURNAL  
 TITLE  
 AUTHORS  
 REFERENCE  
 COMMENT

Center: Sanger Centre  
 Center code: SC  
 Web site: <http://www.sanger.ac.uk>  
 Contact: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)  
 Project Information  
 Center project name: bai37a12  
 ----- Summary Statistics -----  
 Assembly program: XCAP4; version 4.5  
 Sequencing vector: plasmid; L08752; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Consensus quality: 164729 bases at least Q40  
 Consensus quality: 165486 bases at least Q30  
 Consensus quality: 166169 bases at least Q20  
 Insert size: 167063; sum-of-contigs  
 Insert size: 171878; 3.4% error; agarose-fp  
 Quality coverage: 8.28x in Q20 bases; sum-of-contigs quality coverage: 8.11x in Q20 bases; agarose-fp  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently consists of 9 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

* 1 9373: contig of 9373 bp in length
* 9374 9473: gap of 100 bp
* 9474 29891: contig of 20418 bp in length
* 29892 29991: gap of 100 bp
* 29992 36239: contig of 6248 bp in length
* 36240 36339: gap of 100 bp
* 36340 103784: contig of 67445 bp in length
* 103785 103884: gap of 100 bp
* 103885 114878: contig of 10994 bp in length
* 114879 114978: gap of 100 bp
* 114979 126347: contig of 11369 bp in length
* 126348 126447: gap of 100 bp
* 137765 137864: contig of 11317 bp in length
* 137865 163160: contig of 25296 bp in length
* 163161 163260: gap of 100 bp
* 163261 167863: contig of 4603 bp in length.
*
Location/Qualifiers
1. 167863
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-137A12"
/clone_id="RP11-11.1"
1. 9373
/misc_feature
/note="assembly_fragment:02360
fragment_chain:1
clone_end:17
vector_side:left"
9474. 29891
/misc_feature
/note="assembly_fragment:01854
fragment_chain:1"
29992. 36239
/misc_feature
/note="assembly_fragment:00805
fragment_chain:1"
36340. 103784
/misc_feature
/note="assembly_fragment:01467
fragment_chain:1"
103885. 114878
/misc_feature
/note="assembly_fragment:00991
fragment_chain:2"
114979. 126347
/misc_feature
/note="assembly_fragment:01503
fragment_chain:2"
126448. 137764
/misc_feature
/note="assembly_fragment:01297
fragment_chain:2"
137865. 163160
/misc_feature
/note="assembly_fragment:00117
fragment_chain:2"
163261. 167863
/misc_feature
/note="assembly_fragment:01267
fragment_chain:2
clone_end:SP6
vector_side:right"
BASE COUNT 43927 a 39935 c 40002 g 43197 t 802 others
ORIGIN
Query Match 46.7% Score 634.4; DB 2; Length 167863;
Best Local Similarity 99.8%; Pred. No. 2.9e-128;
Matches 635; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 7 GCCATGTCACACAGAGGCTCCGATCCCGAGAGGGCTTGGCAGGCGCTGCTGTGGGA 66
DB 75064 GCCATGTCACACAGAGGCTCCGATCCCGAGAGGGCTTGGCAGGCGCTGCTGTGGGA 75123
OY 67 GCCCAACCTGCTCTCTGCTCCCAATGTGCGGCTCTCTCTGCTGTGTGGCTCTTCTG 126
DB 75124 GCCCAACCTGCTCTCTGCTCCCAATGTGCGGCTCTCTCTGCTGTGTGGCTCTTCTG 75183
OY 127 CTACGTGAGGCTCTCTCTGCTCCACGCGGCGGCGCAGCCCGCTCCCGGGAGAG 186
DB 75184 CTACGTGAGGCTCTCTCTGCTCCACGCGGCGGCGCAGCCCGCTCCCGGGAGAG 75243

```

```

OY 187 GAGATCGTGTTCAGAGAGCTCAACGGCAGCGTCCCTGGCTGGGCGCCCTGCC 246
DB 75244 GAGATCGTGTTCAGAGAGCTCAACGGCAGCGTCCCTGGCTGGGCGCCCTGCC 75303
OY 247 AGCGTGTGTGCTCCCTTCAGAGCGCTTGGGAGAGCGCTGCTACTAGAGCTGAGCAGAC 306
DB 75304 AGCGTGTGTGCTCCCTTCAGAGCGCTTGGGAGAGCGCTGCTACTAGAGCTGAGCAGAC 75363
OY 307 TCCGCTGTGCAAGTGTGAGGGGCTGACAGTCACTGCTGGGCGCAGCGCTGAGCTGTG 366
DB 75364 TCCGCTGTGCAAGTGTGAGGGGCTGACAGTCACTGCTGGGCGCAGCGCTGAGCTGTG 75423
OY 367 GGTGAGCAGAGCGCGGACCTTACCTGTCGACCATCAATGAGATCGGAGCTGGT 426
DB 75424 GGTGAGCAGAGCGCGGACCTTACCTGTCGACCATCAATGAGATCGGAGCTGGT 75483
OY 427 GCATCTGTGCACTGGGATGGGGAGCCCTGTAGCGGTACATATCGGGGGCTGAA 486
DB 75484 GCATCTGTGCACTGGGATGGGGAGCCCTGTAGCGGTACATATCGGGGGCTGAA 75543
OY 487 CTCACCTTCACAGCCCTTGAGGAGAGCACCCTTAATCTGCTGGGGACCTGGGGCTTAC 546
DB 75544 CTCACCTTCACAGCCCTTGAGGAGAGCACCCTTAATCTGCTGGGGACCTGGGGCTTAC 75603
OY 547 ATCTACGCCGGAAGAGTCTGTCCAGCGGTCAAGGTCCTCATGTCACAGCTCCT 606
DB 75604 ATCTACGCCGGAAGAGTCTGTCCAGCGGTCAAGGTCCTCATGTCACAGCTCCT 75663
OY 607 CTGGAAGCCCCAGCCCGCAGACCCCGAAGAGCCCAAG 642
DB 75664 CTGGAAGCCCCAGCCCGCAGACCCCGAAGAGCCCAAG 75699

```

```

RESULT 9
LOCUS AB042271 1251 bp mRNA linear ROD 12-OCT-2000
DEFINITION Rattus norvegicus mRNA for ADAMTS-4, partial cds, clone:ml.
ACCESSION AB042271
VERSION AB042271.1 GI:10799942
KEYWORDS SOURCE

```

```

ORGANISM Rattus norvegicus (strain:Wistar rats) fetus brain cDNA to mRNA,
clone:ml.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (sites)
AUTHORS Satoh, K., Suzuki, N. and Yokota, H.
TITLE ADAMTS-4 (a disintegrin and metalloproteinase with thrombospondin
motifs) is transcriptionally induced in beta-amyloid treated rat
astrocytes
JOURNAL Neurosci. Lett. 289 (3), 177-180 (2000)
MEDLINE 20415831
REFERENCE 2 (bases 1 to 1251)
AUTHORS Satoh, K., Suzuki, N. and Yokota, H.
TITLE Direct Submission
JOURNAL Submitted (27-APR-2000) Kazuki Satoh, Daiichi Pharmaceutical
Co., Ltd., Discovery Research Laboratory, 1-16-13 Kita-ksai,
Edogawa-ku, Tokyo 134-8630, Japan
(E-mail: satohj7@dalichipharm.co.jp, Tel: 81-3-36800151 (ex. 3952),
Fax: 81-3-56964266)

```

```

FEATURES
source 1. 1251
/organism="Rattus norvegicus"
/strain="Wistar rats"
/db_xref="taxon:10116"
/clone="ml"
/tissue_type="brain"
/dev_stage="fetus"
<1..>1251
/codon_start=1
/product="ADAMTS-4"

```

```

/protelin_id="BAB16473.1"
/db_xref="GI:10799943"
/translation="FASTSRFVETLVVADDKMAFHGAGLKHVLLTVMAAAAKFKHP
LTPKPVNLYVTRVILIGSGEVPQGVSAOTKRSPTKQKGNPNDSPDHDVTAI
LFRPDODCVSTSDALCMAGVCTYCDPABSCAIVEDGLOSATFAHDEHVNMLAD
NSKPCANLNGOGSSSRHVMAPVMAHNDPEPWFSCARFTIDELDNCTGCLDKPERA
PLHLPTVPEKRDVADRCQTLTGPDSSHCPLPPLPGLALMCGHLGHAMCOTKHSB
WADOTPCGPACAGGRCGLHVDLKDPIFQAGWGMWGMGDCSRFCGGGVQFSSHD
CTKPVNNGKCYEGRRTPRSCNTKNCPHSALTFREEOCAAYNHRTDLKFSFGPM
DMVPRYGVAPRODCKLTQCARALG"
BASE COUNT      244 a      388 c      347 g      272 t
ORIGIN
Query Match      40.0%: Score 543.4; DB 10; Length 1251;
Best Local Similarity 85.1%: Pred. No. 3.4e-108;
Matches 607; Conservative 0; Mismatches 106; Indels 0; Gaps 0;
OY 646 TTTGCTTACGTAGATAGTATTTGTGGAGACACTGGTGGTGGCAGATGACAAATGGCCCCA 705
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 TTTCCGTTGGTGAATGATTCGTGAGACACTGGTGGTGGCAGATGACAAATGGCACCA 60
OY 706 TTCCACGGTGGGGGCTAAAGCGTACCTGCTAACAGTATGAGCAGCAGCAGCAGCAGC 765
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 TTCCACGGGCGAGGCTTAAAGCACTACTGCTGAGCGTATGCGCGCGCTGGCAAAGCC 120
OY 766 TTCAAGCAGCAGCAGCAGCAGTCTGTCAAGCTTGTGTGTGACTGGCTAGTATCTGT 825
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 TTTAAGCATTCAGACAGCAGCAGCAGTCTGTCAAGCTTGTGTGTGACTGCTAGTATCTGT 180
OY 826 GGGTCAGGCGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTC 885
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 GGGTCAGGCGAGGAGTCCCGCAAGTGGGCGCAAGTGGCGCGCGCGCGCGCGCGCGCTTC 240
OY 886 TGTGCTGCGAGCGGCGCTCAACACCCCTGAGAGACTGCGAGCCTGACACTTTTGACACA 945
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 TGCACCTGGCAGAAAGGCTTAACCCCTCCGAAACGACTCAGATCCGACACTTTGACACG 300
OY 946 GCCATTCTGTTTACCCGTCAGAGCTGTGTGAGTCTCCACTTGGCAGACAGCTGGGTATG 1005
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 GCCATTCTGTTTACCCGTCAGAGCTGTGTGAGTCTCCACTTGGCAGCTGTGGGTATG 360
OY 1006 GCTGATGTGGGCGAGCGCTGTGAGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGC 1065
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 GCTGATGTGGGCGAGCTGTGTGAGCCAGCTAGAGCTGTGCTATTGTGTGAGAGCATGCG 420
OY 1066 CTCCAGTACAGCCTTCACTGCTGCTCATGAAGTGGTCTATGCTTCAACATGCTCCATGAC 1125
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 CTCAGTGTGGCGCTTACCGCGCGCTCATGAAGTGGGCGCATGCTTCAACATGCTCCATGAT 480
OY 1126 AACTGCAAGCCATGATGATTTGAATGGGCTTTGAGACACTCTGCGCATGTCTATGGCC 1185
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 AACTGCAAGCCCTGTGCTAATTTGAATGGGCGAGGAGTTCCTCTGCGCATGTCTATGGCT 540
OY 1186 CTTGTGATGAGCTCATGTGATGAGAGCCCTGTGCTGCGCGCGCGCGCGCGCTTATG 1245
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 541 CTTGTGATGAGCTCATGTGATGAGAGCCCTGTGCTGCGCGCGCGCGCGCGCTTATG 600
OY 1246 ACTGACTTCTGTGACAAATGGCTATGGGCACTGTCTTAGACAAACAGAGGCTCATTTG 1305
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 601 ACTGACTTCTGTGACAAATGGCTATGGGCACTGTCTTAGACAAACAGAGGCTCATTTG 660
OY 1306 CATCTGCGCTGTACTGTGGGAGCTACAGAGCAGTATGACAAAGGCTAGTGTCA 1358
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 661 CATCTGCGAGTGTCTTCTGTGCAAGGACTATGATGTGACGCGCGCTGTGCA 713

```

```

SOURCE
ORGANISM      Mus musculus.
REFERENCE
AUTHORS      Deschamps, S., Gu, W. and Roe, B.A.
TITLE        Mus musculus BAC Clone rp23-395h6
JOURNAL      Unpublished
REFERENCE
AUTHORS      Deschamps, S., Gu, W. and Roe, B.A.
TITLE        Direct Submission
JOURNAL      Submitted (21-NOV-2000) Department Of Chemistry And Biochemistry,
              The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
              OK 73019, USA
REFERENCE
AUTHORS      Deschamps, S., Gu, W. and Roe, B.A.
TITLE        Direct Submission
JOURNAL      Submitted (19-FEB-2002) Department Of Chemistry And Biochemistry,
              The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
              OK 73019, USA
COMMENT
              ----- Genome Center
              Center: Department Of Chemistry And Biochemistry
              The University Of Oklahoma
              Center code: UOKMOR
              -----
FEATURES
source
Location/Qualifiers
1. 211772
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="1"
/clone="rp23-395h6"
/clone_lib="RPI Mouse BAC Library 23"
BASE COUNT      55607 a      48743 c      49203 g      58215 t      4 others
ORIGIN
Query Match      32.0%: Score 434.8; DB 10; Length 211772;
Best Local Similarity 82.8%: Pred. No. 1e-84;
Matches 528; Conservative 0; Mismatches 92; Indels 18; Gaps 2;
OY 8 CCATGTCCGAGACAGGCTGCGATCCCGGAGGCGCTTGGCAGAGCGCTGTGGGAG 67
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 94681 CCATGTCCGAGATGGGCTTGGATCCAGAGAGGCGCTTGGCAGAGCGCTGTGGCAAAT 94622
OY 68 CCAAGCTGCTCTCTGCTCCCATTTGCGGCTCTCTGCTGCTGTGCTGCTGCTGCTGCTG 127
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 94621 TCCAAACCTGCTTGGCGCTTCCACACTGTGC-----AGTGGGAGGCTGC 94577
OY 128 TACTGCTGGCGCTCTCTGCTGCTGAGCGCGGCTGGCCAGCGCCCTCCCGGAGAGG 187
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 94576 TGTGCTGCGCTTCTCTCTGCTTGTAGCGTGCGCGCGCGCGCGCGCGCGCGCGCG 94517
OY 188 AGATGCTTTCCAGAGAGCTCAACGCGC---AGCGTCTGCTGCTGCGGCAACCCCTG 244
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 94516 AGATGCTTTTCCAGAGAGCTCAATGAGCAGTACATCTCTAGTATGACAGCGCTTCTG 94457
OY 245 CCAGGCTTTGTGCGCTTGGACAGCGCTTTGGGAGAGCGCTCTACTAAGCTGAGCAG 304
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 94456 CCAGGCTCTGTACGATTTGGCAGCGCTTTGGGAGATGTTCTACTAAGCTAGAACAG 94397
OY 305 ACTCGGTGTGAGGCTGAGGAGGCTGACAGTACATGAGGCGCGCGCGCGCGCGCTGAGCTGC 364
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 94396 ACCGTGGGCTGAGGTTACAGGCTTTGACTGTACATCTACTGAGGCGCGCGCGCGCGCTGC 94337
OY 365 TGGGTGAGCAGAGCGCTGACACTGAGTGGCAGCATGAATGAGATCCGAGTGG 424
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 94336 TGGGTGGGCGAGAGCGAGTACCTGAGTGGCAGCATGAATGAGATCCGAGTGG 94277
OY 425 TGGCATCTCTGCTGAGTGGAGTGGGAGCGCTGTATAGCGCTGTACAAATATGGGGCGTG 484
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 94276 TGGCATCTCTGCTGAGTGGAGCGGAGCGCTATATGAGGCTACTGAGTACAGTGGGGCGG 94217
OY 485 AACTCAGCTCCAGCGCGCTGGAGGAGGACCGCCCTTAAGTCTGTGCTGGGAGACTGGGCGTC 544

```

```

RESULT 10
AC084821/c 211772 bp DNA linear ROD 19-FEB-2002
LOCUS      AC084821
DEFINITION Mus musculus chromosome 1 clone rp23-395h6, complete sequence.
ACCESSION AC084821
VERSION    AC084821.25 GI:18702388
KEYWORDS   HTG.

```



```

|||||
Db 94216 AACTCACCTCAGCTCTGAGAGAGAGCCCTTAATCTGTCGGGAGACCGGGGCTC 94157
QY 545 ACATCTACACCGGAGAGTCTCTGCGAGGGTCAAGGCCATCGACATCAAGGCTC 604
|||||
Db 94156 ACATCTACACCGGAGAGTCTCTGCGAGGGTCAAGGCCATCGACATCAAGGCTC 94097
QY 605 CTCTTGGAAGCCCGAGCCCGAGACCCCGAGAGCCCAAG 642
|||||
Db 94096 CTCTGGAGAGCCGAGTCCCATTTCCCGCAGACCAAG 94059
|||||

RESULT 11
AC093409 227141 bp DNA linear HTG 23-AUG-2001
LOCUS Mus musculus clone RP23-12J3, WORKING DRAFT SEQUENCE, 14 unordered
DEFINITION pieces.
ACCESSION AC093409 GI:15281247
VERSION AC093409.1
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 227141)
AUTHORS Martin,J., Hosseini,R., Peng,Y., Peng,Z., Rubin,E.M. and Cheng,J.F.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 227141)
AUTHORS Martin,J., Hosseini,R., Peng,Y., Peng,Z., Rubin,E.M. and Cheng,J.F.
TITLE Direct Submission
JOURNAL Submitted (23-AUG-2001) Berkeley PGA, Lawrence Berkeley National
Laboratory, 1 Cyclotron rd., Berkeley, CA 94720, United States of
America
COMMENT Draft Sequence Produced by Berkeley PGA
Pga.tbl.gov
Center Code: PGABERK.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 2601: contig of 2601 bp in length
* 2602 2701: gap of unknown length
* 2702 4917: contig of 2216 bp in length
* 4918 5017: gap of unknown length
* 5018 11010: contig of 5993 bp in length
* 11011 11110: gap of unknown length
* 11111 20299: contig of 9189 bp in length
* 20300 20399: gap of unknown length
* 20400 26860: contig of 6461 bp in length
* 26861 26960: gap of unknown length
* 26961 38473: contig of 11513 bp in length
* 38474 38573: gap of unknown length
* 38574 45940: contig of 7367 bp in length
* 45941 46040: gap of unknown length
* 46041 58468: contig of 12428 bp in length
* 58469 58568: gap of unknown length
* 58569 75898: contig of 17330 bp in length
* 75899 75998: gap of unknown length
* 75999 86499: contig of 13501 bp in length
* 86500 88599: gap of unknown length
* 88600 99646: contig of 10047 bp in length
* 99647 99746: gap of unknown length
* 99747 118586: contig of 18840 bp in length
* 118587 118686: gap of unknown length
* 118687 150303: contig of 31617 bp in length
* 150304 150403: gap of unknown length
* 150404 227141: contig of 76738 bp in length.

```

```

FEATURES
source location/Qualifiers
1..227141
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP23-12J3"

BASE COUNT 59478 a 55126 c 53628 g 57553 t 1356 others
ORIGIN

Query Match 32.0%: Score 434.8; DB 2; Length 227141;
Best Local Similarity 82.8%: Pred. No. 1e-84;
Matches 528; Conservative 0; Mismatches 92; Indels 18; Gaps 2;

QY 8 CCATGTCCACAGACGCTCGCATCCCGAGAGGGCTTGACAGAGGCGCTGTGGGAG 67
Db 137430 CCATGTCCACAGATGGCGTTTGATCCAGAGAGGGGTACTGAGGCACTGGCAAAAT 137489
QY 68 CCACACCTGCTCTCTGCTCCCATTTGCGCGCTCTCTGCTGTGTGCTGCTTCTGC 127
Db 137490 TCCACACCTGCTGCTGCGCTTCCACACTGTGC-----AGTGGCGGAGGCTGC 137534
QY 128 TACGTGCTGCTCTCTCTGCTGCGCGCTGCGCGGCGGCGGCGGCGGAGG 187
Db 137535 TGCTGCTGCGCTTCTCTCTGCTGCTGCTGCGCGGCGGCGGCGGAGG 137594
QY 188 AGATCGTGTTCACAGAGAGCTCAACGCGC---AGCGTCTGCTGCTGCTGCGGCAACCCCTG 244
Db 137595 AGATCGTGTTCACAGAGAGCTCAATGCGATGAGCATCTACTGATGATCAGGCGTTCCTG 137654
QY 245 CCAGGCTGTGTGCGCGCTTGACAGGCGCTTTGGGAGAGACGCTGCTACTAGAGCTGACGACG 304
Db 137655 CCAGGCTGCTGTACCGATTCGACGCGCTTTGGGAGAGATCTTCTACTAGACTAGAACG 137714
QY 305 ACTCGGTGTGAGGTGCGAGGCGCTGACAGTGCATGCTACTGCGGCGGCGGCTGAGCTGC 364
Db 137715 ACCTGGGGGTGAGGTAGAGGCTTTGACTGTACACTACTGCGGCGGCGGAGATGC 137774
QY 365 TGGGTGAGAGAGGCTGCGACACTACTGCTGCGGCGGCGGCGGCGGAGTGC 424
Db 137775 TGGGTGAGGAGAGGCGAGTACTACTGCTGCGGCGGCGGCGGAGATGC 137834
QY 425 TGGCATCTCTGCTGAGGATGAGGAGGCGGCTTTGAGGCTGTACAAATATGCGGCGCTG 484
Db 137835 TGGCATCTCTGCTGAGGAGGAGGCGGCGGCTTTGAGGCTGTACAAATATGCGGCGCTG 137894
QY 485 AACTCACCCTCGAGCGGCTTGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCTG 544
Db 137895 AACTCACCCTCGAGCGGCTTGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCTG 137954
QY 545 ACATCTACACCGGAGAGTCTCTGCGAGGGTCAAGGCCATCGACATCAAGGCTC 604
Db 137955 ACATCTACACCGGAGAGTCTCTGCGAGGGTCAAGGCCATCGACATCAAGGCTC 138014
QY 605 CTCTTGGAAGCCCGAGCCCGAGACCCCGAGAGCCCAAG 642
Db 138015 CTCTTGGAAGCCCGAGTCCCATTTCCCGCAGAACCAAG 138052

RESULT 12
AC099236/c 84868 bp DNA linear HTG 13-JUL-2002
LOCUS Rattus norvegicus clone CH230-99J19, *** SEQUENCING IN PROGRESS
DEFINITION **, 46 unordered pieces.
ACCESSION AC099236
VERSION AC099236.4 GI:21735644
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 84868)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alibrooks,S.L., Amaralunge,H.C., Are,J.R., Ayale,M., Banks,T.,

```

Barbarta, J., Benton, J., Blmage, K., Blankenburg, K., Bonnin, D.,  
Bouck, J., Bowle, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P.,  
Bunay, C., Burck, P., Burkett, C., Burrell, K.L., Byrd, N.C.,  
Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,  
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,  
Daviya, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,  
DeJarey, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,  
Douthett, C., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,  
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
Gabriel, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,  
Gorelli, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,  
Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,  
Honsi, F., Howard, S., Huber, J., Huliy, S., Hume, J., Jackson, L.E.,  
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,  
Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,  
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louised, H.,  
Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,  
Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mel, G., Metzger, M.,  
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,  
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogun, M., Okunolu, G.,  
Ongunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,  
Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,  
Scherer, S., Scott, G., Shen, H., Shoshitani, N., Slason, I.,  
Sodergren, E., Sonake, T., Sparks, A., Stanley, H., Stone, H.,  
Sutton, A., Sytek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,  
Umanil, K., Vasquez, L., Vera, V., Villalón, D., Vinsón, R., Wang, Q.,  
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,  
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
Wainstock, G., and Gtdbs, R.

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

Submitted (09-NOV-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 84868)  
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jul 12, 2002 this sequence version replaced gi:17973412.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: CHW  
Center clone name: CH230-99J19  
----- Summary Statistics  
Sequencing vector: Plasmid  
Chemistry: Dye-terminator Big Dye 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 37772 bases at least Q40  
Consensus quality: 40463 bases at least Q30  
Consensus quality: 42506 bases at least Q20

\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank/draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 46 contigs. The true order of the pieces

\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1414: contig of 1414 bp in length  
1415 1514: gap of unknown length  
1515 2536: contig of 1022 bp in length  
2537 2636: gap of unknown length  
2637 4007: contig of 1371 bp in length  
4008 4107: gap of unknown length  
4108 5478: gap of 1370 bp in length  
5478 5577: gap of unknown length  
5578 6645: contig of 1068 bp in length  
6646 6746: gap of unknown length  
6746 8297: contig of 1551 bp in length  
8297 8396: gap of unknown length  
8396 10028: contig of 1632 bp in length  
10028 10128: gap of unknown length  
10128 11443: contig of 1315 bp in length  
11443 11543: gap of unknown length  
11543 12949: contig of 1406 bp in length  
12949 13049: gap of unknown length  
13049 14321: contig of 1272 bp in length  
14321 14421: gap of unknown length  
14421 15523: contig of 1102 bp in length  
15523 15623: gap of unknown length  
15623 16796: contig of 1173 bp in length  
16796 16896: gap of unknown length  
16896 18584: contig of 1688 bp in length  
18584 18684: gap of unknown length  
18684 20481: contig of 1797 bp in length  
20481 22037: contig of 1456 bp in length  
22037 22137: gap of unknown length  
22137 23464: contig of 1327 bp in length  
23464 23564: gap of unknown length  
23564 24657: contig of 1093 bp in length  
24657 24757: gap of unknown length  
24757 26276: contig of 1519 bp in length  
26276 26377: gap of unknown length  
26377 27440: contig of 1064 bp in length  
27440 27540: gap of unknown length  
27540 29128: contig of 1588 bp in length  
29128 29228: gap of unknown length  
29228 31090: contig of 1862 bp in length  
31090 31190: gap of unknown length  
31190 32310: contig of 1120 bp in length  
32310 32410: gap of unknown length  
32410 33935: contig of 1525 bp in length  
33935 34035: gap of unknown length  
34035 37032: contig of 2997 bp in length  
37032 37132: gap of unknown length  
37132 38855: contig of 1723 bp in length  
38855 38955: gap of unknown length  
38955 41393: contig of 2438 bp in length  
41393 41493: gap of unknown length  
41493 43503: contig of 2010 bp in length  
43503 43603: gap of unknown length  
43603 45019: contig of 1416 bp in length  
45019 45119: gap of unknown length  
45119 47142: contig of 2023 bp in length  
47142 47242: gap of unknown length  
47242 49211: contig of 1969 bp in length  
49211 49311: gap of unknown length  
49311 50680: contig of 1369 bp in length  
50680 50780: gap of unknown length  
50780 52793: contig of 2013 bp in length  
52793 52893: gap of unknown length  
52893 55142: contig of 2149 bp in length  
55142 56319: gap of unknown length  
56319 56319: contig of 1177 bp in length

Query Match	Best Local Similarity	Matches	523: Conservative	29.2%: Score 397.2; DB 2; Length 84868; Pred. No. 1.8e-76;	0: Mismatches 108;	Indels 11;	Gaps
56320	56419: gap of unknown length						
56420	57872: contig of 1453 bp in length						
57873	57972: gap of unknown length						
57973	61529: contig of 357 bp in length						
61530	61629: gap of unknown length						
61630	63596: contig of 1967 bp in length						
63597	63696: gap of unknown length						
63697	65966: contig of 2270 bp in length						
65967	66067: gap of unknown length						
66067	68070: contig of 2004 bp in length						
68071	68170: gap of unknown length						
68171	69722: contig of 1552 bp in length						
69723	69822: gap of unknown length						
69823	72950: contig of 3128 bp in length						
72951	73050: gap of unknown length						
73051	74867: contig of 1917 bp in length						
74968	75067: gap of unknown length						
75068	77568: contig of 2501 bp in length						
77569	77668: gap of unknown length						
77669	80667: contig of 2999 bp in length						
80668	80767: gap of unknown length						
80768	82926: contig of 2159 bp in length						
82927	83026: gap of unknown length						
83027	84868: contig of 1842 bp in length.						
83027	Location/Qualifiers						
1. 84868							
/organism="Rattus norvegicus"							
/db_xref="taxon:10116"							
/clone="CH230-99319"							
BASE COUNT	20641 a	18246 c	17435 g	20577 t	7969 others		
ORIGIN							
Query Match	29.2%: Score 397.2; DB 2; Length 84868;						
Best Local Similarity	81.5%: Pred. No. 1.8e-76;						
Matches	523: Conservative	0: Mismatches	108;	Indels	11;	Gaps	
8	CCATGCCAGACAGGCTGCGATCCCCGGAGGGCGTGTGGACGGCGCTGTGTGGGAG	67					
25604	CCATGTCACAGGATGGGCTGTGATCCACGAGGGGCTGTGATGGGACACAGCTCCGAGAC	25545					
68	CCCAACCCCTGCCTCCGTGCTCCCATTTGTCGCCCTCTCCCTGGGTG---TGTGGCTGTT	123					
25544	TCCAACCCCTGCCCTGGCGCCCGCCACTGTGCCCTTTCTCGGGTTTGGAGCGTTGGCGACAG	25485					
124	CTGTACTGTGTGGCTCTCTCTGCCCTGAGCCCGGTGGCCAGCCCTCCCGGGGAG	183					
25484	CTGTCTGTGTGTCTCCCTC-CTCCTGCTTTAAGCCCTG--CGCCAGGCCCTCCCTCCACAG	25428					
184	GAGGAGATGTTTTCACAGAGAGCTCAAGGC---AGCTCCTGCTGGCTGGCGGACC	240					
25427	GAGGAGATCG-GGTTCACAGAGAGCTCAAGGCAGTAGTATATATCTGTGAGGCGTC	25369					
241	CCGTGCAGGCTTTGTGCGCTTGCAGGCGCTTTGGGAGAGCGCTGTACTAGAGCTGAG	300					
25368	CCGTGCAGGCTGTGTACCGATTACAGCGCTTTGGGAGAGATATTGTACTAGAACTAGAG	25309					
301	CAGGACTCCGGTGTGAGGTTCGAGGCGCTGACAGTCACTACTACTGGGCCAGCGCTGAG	360					
25308	CAGGATCCGGGGGTGAGGTAGAGGGGCTGACAGTCACTACTACTGGGCCAGCGCTTGA	25249					
361	CTGTGTGGTGGAGCAGAGGCTGGCACTTACCTGACTGGGACCATCAATGAGATCCGAG	420					
25248	ATGTGGGGTGGGGAGAGGCAAGTACCTACCTGATCGGACCATCAATGAGATCCGAG	25189					
421	TGGGTGGCATCTCTGTGACTGTGGATGGGGAGCCCTTTAGCGGTGTACAAATATCGGGG	480					
25188	TGGGTGGCATCTCTGTGACTGTGGATGGGGAGCCCTTTAGCGGTGTACAAATATCGGGG	25129					
481	GTGGAATCTCACTCTCAAGCCCTTGAGGAGGACCCCTTAATCTGCTGGGGACCTGGG	540					
25128	GCGGAATCTCACTCTCAAGCCCTTGAGGAGGACCCCTTAATCTGCTGGGGACCTGGG	25069					
541	GCTCAATCTTACGCGGAGAGATCTTGCACGCTCAAGGTTCATGTGCAACGTCAAG	600					

Df	25068	GCTCAGATTCCTACGCCCGGAAAGAGTCTCCACCAACCAGAGGTCCCATATGCACAAGCTCAAG	25009
Qy	601	GCTCCTTTGAAGCCCCAGACCCGAGAGAGCCAG	642
Df	25008	GCTCCTTGCGGAGCCCCAGTCCCATTTTCCCGAGAACCAAG	24967
 RESULT 13 AF368321			
LOCUS	AF368321	576 bp	mRNA linear
DEFINITION	Equus caballus aggreccanase-1 mRNA, partial cds.		MAM 15-MAY-2001
ACCESSION	AF368321		
VERSION	AF368321.1	GI:14039920	
KEYWORDS			
SOURCE			
ORGANISM	Equus caballus. Equus caballus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Perissodactyla; Equidae; Equus.		
AUTHORS	Plannery,C.R. and Little,C.B.		
TITLE	Expression and activity of equine aggreccanases		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 576)		
AUTHORS	Plannery,C.R. and Little,C.B.		
TITLE	Direct Submission		
JOURNAL	Submitted (05-Apr-2001) School of Biosciences, Cardiff University, Museum Avenue, Cardiff CF10 3US, United Kingdom		
FEATURES			
source	location/Qualifiers		
	1..576		
	/organism="Equus caballus"		
	/db_xref="taxon:9796"		
	/cell-type="articular cartilage chondrocyte"		
	<1..>576		
	/function="Metalloproteinase"		
	/note="ADAMTS family member: similar to Homo sapiens aggreccanase-1 encoded by Genbank Accession Number AF148211 and Bos laurus aggreccanase-1 encoded by Genbank Accession Number AF192770"		
	/codon_start=1		
	/product="aggreccanase-1"		
	/protein_id="AKK53425.1"		
	/db_xref="GI:14039921."		
	/translation="MADVGVTCDDPARSCAIVEDDGLASFTPAHELGHVFMKLHDSN PCVGINGPGSTSRHYMAVMHVDPEEWSFCSAPFTIDPNDNGRCLDKPEARPLR LPVPFGKDYADRQCCLDTFGDRSHRCMLDPSPCALWCSGLNHLMACQTKNSPMAL GPSGPDAOCMGRCILHMDLOEFNFIPOLAGW"		
BASE COUNT	99 a 181 c 160 g 136 t		
ORIGIN			
Query Match	21.3%; Score 288.8; DB 4; Length 576;		
Best Local Similarity	90.6%; Pred. No. 1.3e-52;		
Matches	308; Conservative 0; Mismatches 32; Indels 0; Gaps 0;		
Qy	1003	ATGGCGATGTGGGACCGCTGTGTGACCGCGGCTGGAGCTGTGGCAATTGTGGAGATGAT	1062
Df	1	ATGGCTGATGTGGGACTGTGTGTGTGACCCGGCTCGGAGCTGTGTAATGTGGAGATGAT	60
Qy	1063	GGGCTCCAGTCAGCCTTACTGCTGTCTGTGTAAGTGGGTTCATAATGCTTCAT	1122
Df	61	GGGCTCCAGTCAGCCTTACTGCTGTCTGTGTAAGTGGGTTCATAATGCTTCAT	120
Qy	1123	GACAACTCCAAGCCATGCATTCAGTTTGATGGGCTTTTAGAGACCTCTGGCCATGTCAAG	1182
Df	121	GACAAATTCAAAACCAATGTGTGTGTGATGGGCTCGGAGACCACTCCCGCATGTCAAG	180
Qy	1183	GGCCCTGTATGGCTCATGTGATCCCTGTAGGAGACCCCTGTCCCCTGCAGTCCGCTTC	1242
Df	181	GGCTCGTATGGCTCAGTGTGATCCCGAGAGACCCCTGTCTCCCTGCAGTCTCGCTTC	240
Qy	1243	ATCACTGACTTCTGTGACATGGCTATGGGCACTGTCTTTAGACAAACGAGAGCTCCA	1302
Df	241	ATCACTGACTTCTGTGACATGGCTATGGGCACTGTCTTTAGACAAACGAGAGCTCCC	300



```
PC C12N15/09.C07K16/40.C12N1/15.C12N1/19.C12N1/21.C12N5/10, PC
C12N9/50.C12Q1/37,
PC C12N15/00.C12N5/00
CC
EH Key Location/Qualifiers
FT source 1..2853
FT Location/Qualifiers
1..2853
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 704 a 734 c 816 g 599 t
ORIGIN
Query Match 20.6%; Score 280.2; DB 6; Length 2853;
Best Local Similarity 62.3%; Pred. No. 8.3e-51;
Matches 457; Conservative 0; Mismatches 273; Indels 3; Gaps 1;
OY 626 GACCCGAGAGCCCAAGCGCTTGTCTCACTGAGTAGATTTGTGAGACACTGTGTGTG 685
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 686 GAAAGCATAGAAAGAGGATTTGTCTCAGTCACCGCTATGTGAAACCATGCTGTGTG 745
OY 686 CAGATGACAGATGGCCGATTCACGGGCGGGCTAAAGCGCTAACCTGCTAACAGTGA 745
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 746 CAGACCATGTCGATGGCAGATTTCACGGCAGTGTCTTAAGCATTACCTTTCACGTTGT 805
OY 746 TGGCAGCAGACCCCAAGGCTTCAAGCACCAGCATCCGCAATCCTGTACGCTGTGTG 805
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 806 TTTCGTGGCAGCCAGATTGTACAAACACCCGACATTCGTATTACGTTACGCTGTGTG 865
OY 806 TGACTCGGCTAGTATCTCTGGGGTCAGGGGAGAGGGGCCCAAGTGGGGCCAGTCTG 865
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 866 TGGTGAAGATCTTGGTCAATCCAGATGAACAGAGGGGCCGAGAGTGAACCTCAATGCTG 925
OY 866 CCCAGACCTGGCGAGCTTCTGTGGCTGGCAGCGGGGCTCAACACCCCTGAGGACTGG 925
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 926 CCTTCACCTGTGGAGACTTTTGCAACTGGCAGAGAGACACACCCACTGACCGGG 985
OY 926 ACCCTGACCATTGTGACACAGCCATTCTGTTCACCGTCAGAGACCTGTGTGAGTCTCA 985
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 986 ATGCAGAGACATATGACACAGCAATTTCTTACACAGACAGSACTTGTGTGTCCAGTA 1045
OY 986 CTTCGACACAGCTGGGTATGTGTGATGTGGGACACCGTCTGTGACCCGGCTCGAGCTGTG 1045
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1046 CATGTGATACCTTGTGGGATGCTGATGTGGTGAATGCTGTGATCGACAGAGAGCTGCT 1105
OY 1046 CCATTGTGAGATGTGGGGCTCCAGTCAGCCTTCATGCTGCTCATGTAACAGTGGTCAATG 1105
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1106 CCGTCATAGAGATATGTGTTTACAACTGCTTCACACAGCCCATGTAATTAGGCCACG 1165
OY 1106 TCTTCACATGCTCCATGACAACTCCAAAGCATGATGATTTGAATGGGCTTTGAGCA 1165
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1166 TGTTCACATGTCACATATGATGATGCAAAAGCATGTGCCACCTTAATG---TGTAAAC 1222
OY 1166 CCTTCGCGCATGTCAATGGCCCTGTGATGGCTCAATGTGGATCCTTGAGAGCCCTGTGCC 1225
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1223 AGGATTCACACATGTGATGGGCTCAATGCTTCCAACTGTGACCACAGCCAGCTTGGTCTC 1282
OY 1226 CTTGAGTGGCCGCTTCACTCACTGACTTCCGACAAATGGGCAATGGGCACTGTCTTATG 1285
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1283 CTTGCAATGCTTACATGATTTATCATATTTCTGTGATATGCTCATGTGGGAATGTTGATGG 1342
OY 1286 ACAAAACAGAGGCTCCATTGCTATGCTGTGATCGGAGCTACAGAGACAGATGACA 1345
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1343 ACAAGCTTCAGAAATCCCATACAGCTCCAGGGGATCTCCCTGGCACTGTGATGATGCCA 1402
OY 1346 AGGGGTAGCTGA 1358
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1403 ACCGGCAGTGCCA 1415
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 14:40:26 ; Search time 58.6837 Seconds  
(without alignments)  
1727.484 Million cell updates/sec

Title: US-10-050-200-9  
Perfect score: 2614  
Sequence: 1 MLLQWASLLLCARLPLAAV.....LDLPRKQILGQYKDDDDKG 492

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL.21.\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mmc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_proteol:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriaph:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	770.5	29.5	950	4	Q8TE58
2	548	21.0	269	6	Q9GL54
3	493	18.9	2165	5	Q19791
4	464	17.8	1095	4	Q8TE56
5	461.5	17.7	1072	4	Q8TE57
6	443	16.9	1054	5	Q9W493
7	420	16.1	1207	4	Q8TE59
8	418	16.0	1223	4	Q8WXS8
9	418	16.0	1223	4	Q8TE55
10	410.5	15.7	1159	4	Q8TE58
11	396.5	15.2	1688	5	Q8SXB0
12	392	15.0	1081	4	Q8TE60
13	367.5	14.1	1229	5	Q9VFE1
14	354	13.5	187	6	Q95N23
15	339	13.0	899	13	Q8UVP1
16	327.5	12.5	914	13	Q12960

17	306	11.7	873	13	Q42595	042595 xenopus lae
18	297	11.4	922	13	Q8UVP2	Q8UVP2 coturnix co
19	296.5	11.3	845	11	Q61072	Q61072 mus musculu
20	294.5	11.3	192	6	Q95N24	Q95N24 equus caball
21	291	11.1	1091	5	Q9W126	Q9W126 drosophila
22	281.5	10.8	791	11	Q8R533	Q8R533 mus musculu
23	279.5	10.7	806	11	Q8R534	Q8R534 mus musculu
24	279.5	10.7	1427	4	Q96L37	Q96L37 homo sapien
25	276.5	10.6	825	11	Q8R3D3	Q8R3D3 mus musculu
26	275.5	10.5	819	4	Q13443	Q13443 homo sapien
27	264.5	10.1	505	13	Q73795	Q73795 agkistrodon
28	263	10.1	814	4	Q96C78	Q96C78 homo sapien
29	262.5	10.0	617	13	Q90499	Q90499 echis pyram
30	261.5	10.0	483	13	Q91AB0	Q91AB0 agkistrodon
31	260.5	10.0	610	13	Q9Y120	Q9Y120 agkistrodon
32	258	9.9	1023	5	Q967H9	Q967H9 strongyloce
33	257	9.8	466	13	Q91A66	Q91A66 agkistrodon
34	252.5	9.7	604	13	Q9P748	Q9P748 atracaspis
35	252	9.6	411	13	Q92031	Q92031 agkistrodon
36	251	9.6	482	13	Q9PVK9	Q9PVK9 agkistrodon
37	250.5	9.6	400	13	Q91AY1	Q91AY1 agkistrodon
38	250	9.6	479	13	Q9PWJ0	Q9PWJ0 agkistrodon
39	249.5	9.5	400	13	Q91AV3	Q91AV3 agkistrodon
40	249.5	9.5	610	13	Q93523	Q93523 bothrops ja
41	249.5	9.5	610	13	Q80C88	Q80C88 bothrops in
42	249.5	9.5	612	13	Q90Z13	Q90Z13 trimeresuru
43	247	9.4	409	13	Q80C89	Q80C89 bothrops in
44	247	9.4	466	13	Q91AX7	Q91AX7 agkistrodon
45	247	9.4	478	13	Q98995	Q98995 vipera lebe

## ALIGNMENTS

## RESULT 1

Q8TE58 PRELIMINARY: PRT: 950 AA.  
ID Q8TE58;  
AC Q8TE58;  
DT 01-JUN-2002 (TREMBLrel. 21, Created)  
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Metalloprotease disintegrin 15 with thrombospondin domains.  
DE ADAMTS15.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21856482; PubMed=11867212;  
RA Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V.,  
RA Lopez-Otin C.,  
RT Cloning, expression analysis, and structural characterization of  
RT seven novel human ADAMTS, a family of metalloproteinases with  
RT disintegrin and thrombospondin-1 domains.";  
RL Gene 283:49-62(2002).  
DR EMBL: AJ315733; CAC86014.1; -  
KW Integrin; Protease.  
SQ SEQUENCE 950 AA: 103286 MW: 50FBE18285CCCC3B CRC64;

Query Match 29.5%; Score 770.5; DB 4; Length 950;

Best Local Similarity 42.8%; Pred. No. 5, 1e-54;  
Matches 180; Conservative 49; Mismatches 133; Indels 59; Gaps 12;

QY 82 SSGKVGKYLIVAGRRFLDLERDGSVGIAGF-----VPAGC--GTSAPMHRSHCPFR 133  
DB 48 SDDGGIFQITFAFOEDFYHLTPDAQFLAPAFSTEHLGVPLGLTGSSDLR---RCFYS 104  
QY 134 GYVDSRSLAVFDLCGGIDGFFAVKHARYTLKPLIRGWMAEENKGRYVGDSSARILAHY 193  
DB 105 GYVNAEPDSFAVSLCGGIRGAPGYRGAETVISP---PNAAPAAQNRNSQA---HLL 157  
QY 194 TREGSFSEALPPRASCETPASTPEAHEHAPAHNSNPGRAALAS---QLDDQSALSP--- 246

```

DB 158 QRRG-----VPGSPS-----GDPTRSGVAGSNWPAAL--RALDPYKP 193
QY 247 -AGSGPQTMRRRRR-----ISRAVOELLVLVADAMARLYGGLHYLLTLASIN 299
DB 194 RAAGGCS---ASRRSGCAKRFVSIPIRYVELLVADDESMVAFHGDLCYLLTLATA 250
QY 300 RLYSHASIEHRIALVAVKVVVLGDKDSLVSNAATLTKNCKMOHQHQLGDDHEEH 359
DB 251 RLYRHSPIINPIVYVYKLLLRDRSGPKVGNALTLRNCAMQKLNKYSDKRPEXW 310
QY 360 DAAILFTREDLCGHSCDTLGMADVGTICSPERSCAVIDDGLMAAFYVAHEIGLGLS 419
DB 311 DTAILEFTRODLGATTCDTLGMADVGTICDPKRSCEVIDDGLPSAFYVAHEIGLVENP 370
QY 420 HDSKRCETFCSTEDKRLMSITLSDASKPMSKCTSAITEFLDDGNCILLDPRQ 479
DB 371 HDNVKCEVFGKLRANHMSPFLIQRANPMSACSAAITFDLDSGHGCLLDQSPK 430
QY 480 I 480
DB 431 I 431

```

RESULT 2  
ID Q9GL54 PRELIMINARY: PRT: 269 AA.

```

AC 09GL54:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE 01-UN-2002 (TREMBLrel. 21, Last annotation update)
DE Aggrecanase-2 (Fragment).
GN ADAMTS-11.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxId:9986;
RM 11
RP SEQUENCE FROM N.A.
RA Goad D.L., Goad M.E.;
RT "Molecular cloning of lapine aggrecanase-2 (ADAMTS-11) from articular
RT chondrocytes."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF317415; AAG33062.1;
DR HSSP: Q9PW35; 1BUD.
DR MEROPS: M12.225;
DR InterPro: IPR001590; Reprolysin.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR000130; Zn_Mpeptidse.
DR Pfam: PF01421; Reprolysin; 1.
DR Pfam: PF00090; Tsp_1; 1.
DR SMART: SM00209; TSP1; 1.
DR PROSITE: PS50215; ADAM_MERPO; 1.
DR PROSITE: PS50092; TSP1; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 269
SQ SEQUENCE 269 AA; 29193 MW; 97A1CA80B33452FA CRC64;

```

Query Match 21.0%; Score 548; DB 6; Length 269;  
Best Local Similarity 98.1%; Pred. No. 1.2e-36;  
Matches 103; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 378 TLGMADVGTICSPERSCAVIDDGLMAAFYVAHEIGLGLSHDSDKFCETFGSTEDK 437
DB 1 TLGMADVGTICSPERSCAVIDDGLMAAFYVAHEIGLGLSHDSDKFCETFGSTEDK 60
QY 438 LMSILTSIDASKPMSKCTSAITEFLDDGNCILLDPRQILG 482
DB 61 LMSILTSIDASKPMSKCTSAITEFLDDGNCILLDPRQILG 105

```

RESULT 3

```

Q19791
ID Q19791 PRELIMINARY: PRT: 2165 AA.
AC Q19791: 027524;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 08, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE F25H8.3 protein.
GN F25H8.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxId:6239;
RM 11
RP SEQUENCE FROM N.A.
RA Gadsdley S.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RN 12

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE-94150718; PubMed-7906398;
RA Wilson R., Alnscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Keishaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkhen R.,
RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Spoat J., Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38(1994).
DR EMBL: 269361; CA93288.1;
DR EMBL: 269360; CA93288.1; JOINED.
DR EMBL: 269360; CA93287.1;
DR EMBL: 269361; CA93287.1; JOINED.
DR HSSP: P15167; 1DTH.
DR MEROPS: M12.135;
DR InterPro: IPR001590; Reprolysin.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR000130; Zn_Mpeptidse.
DR Pfam: PF01421; Reprolysin; 1.
DR Pfam: PF00090; Tsp_1; 14.
DR SMART: SM00209; TSP1; 18.
DR PROSITE: PS50215; ADAM_MERPO; 1.
DR PROSITE: PS50092; TSP1; 6.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 2165 AA; 244397 MW; FCC3DAB8AA9C4888 CRC64;

```

Query Match 18.9%; Score 493; DB 5; Length 2165;  
Best Local Similarity 32.1%; Pred. No. 6.7e-31;  
Matches 126; Conservative 66; Mismatches 124; Indels 76; Gaps 15;

```

QY 118 GGTSAAPWHRHRSICFYRGVDSGSPRLAVFDLC--GGLDGFPAVKHARTLKLRLGPA 174
DB 160 GGTSAAPWHRHRSICFYRGVDSGSPRLAVFDLC--GGLDGFPAVKHARTLKLRLGPA 215
QY 175 EEKKGKRVGDSGARILHYVTRGCFSEALPPRASCTPASTPEAHENAHNSPGRAL 234
DB 216 EH-----DGASRHRHOLVRK--FDPW-----HKSPDHLN----- 243
QY 235 ASQLDQSLSPAGSGGPQW--N-----RRRRSISRAVOELLVLVADAMARLY 283
DB 244 -STVNETFTTVA-----TWDDQMEDYIERKARSRANSMDHYVELVADTKMYEH 296
QY 284 GGLGOLHYLLTLASINRLYSIASIENHRIALVAVKVVVLGDKDSLVSNAATLTKNPK 343
DB 297 GRLSDYVLTFTPSVAISIRHOSLRASINVVVKLVLTENACGRITONAOOTLDDPCR 356
QY 344 WOHONOLGDDHIEHYDAILFTREDLC-GHNSCDTLGMADVGTICSPERSCAVIDDGL 402
DB 357 WOOYNDPDDSSVOHHDVAIILTRDLCRSGCKCPTGLAELGTWCMOKSCAILEDNGL 416

```







```

Db 212 PPAPDACCFTYGAVALRHGSLASTGCGLMCFIQLNEDFIETPL-----259
QY 181 VYDGSARILHYVTRGFSFELPPRASCETPASTPEAHEHAHNSPGRALASOLL 240
Db 260 -----NDTAITGHPRIYVRKQKSMEEKTEKSAHSHICG--IISDKGR 303
QY 241 QSALSPAGSGSPQWMMRRRSISRAROVELLVADASMARLYGR-GLQHYLLTLASIA 299
Db 304 RSRKIAESGRC-----KRSYKLPQEVNIETVYVADPAMVSYHADARFILLTNMVF 358
QY 300 RLXSHASIEHNIIRLAVVAVVVLGDKDSLEVSNAATTLKFNCKQHONQDGDHDEH- 358
Db 359 NLPRKSLGVQVNMRLVILLLHETPELLYIGHGEKMLSECFKQWHE--EFGKNDIHL 416
QY 359 -----YDAILFTREDLQGH--HSCDTLGMADVGTICSPERSCAVIEDDGLHA 404
Db 417 EMSTWGGEDMTSYDAAILITRKDEPCVHDEPCDTYGIATLSMCKEKKCIATBONGI 476
QY 405 APYVAHEIGHLGLSHD--DSKFCERTFGSTEDKRLMSILTSIDASKPWSKCTSAITTEF 463
Db 477 AFTLAHEMGNHNGIINHNDHPSCADGLHMSGEMIKGNLGDVS---WSRCKEDLERF 532
QY 464 LDDGHCNCLL 473
Db 533 LRSKASNCLL 542

```

## RESULT 8

```

O8WXS8 PRELIMINARY; PRT: 1223 AA.
AC O8WXS8;

```

```

DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE A disintegrin-like and metalloprotease with thrombospondin type 1
  motif 14 precursor.
GN ADAMTS14.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638061; PubMed=11779638;
RA Bolz H., Ramirez A., von Brederlow B., Kubisch C.;
  "Characterization of ADAMTS14, a novel member of the ADAMTS
  metalloproteinase family.";
  EMBL; AF358666; AAL40229.1;
  InterPro: IPR002870; Pep_M12B_propep.
  InterPro: IPR001590; Reprolysin.
  InterPro: IPR000884; TSP1.
DR Pfam; PF01562; Pep_M12B_propep. 1.
DR Pfam; PF01421; Reprolysin. 1.
DR Pfam; PF00090; TSP1. 4.
DR SMART; SM00209; TSP1. 4.
DR PROSITE; PS50215; ADAM_MPRO; 1.
DR PROSITE; PS50092; TSP1. 1.
KW Signal; Integrin; Protease; Metalloprotease.
FT SIGNAL 1 22 POTENTIAL.
SQ SEQUENCE 1223 AA; 133871 MW; 36394ACAD92F170F CRC64;

```

```

Query Match 16.0%; Score 418; DB 4; Length 1223;
Best Local Similarity 27.6%; Pred. No. 3.9e-25;
Matches 133; Conservative 67; Mismatches 176; Indels 106; Gaps 20;

```

```

QY 21 GPAATPAQDK-AGOPTAAAAQPRRQGEVQEAPEPGHPHPLAQRSSKGLVONIDQ 79
Db 59 GPAASAGSMVVDTPTL-----PRHSHLRV---ARSLHP-----92
QY 80 LYSGG-----GKVG-----YLYVAGRRFLDLERDGSVGIACFVAGGTSAPMWH-----126
Db 93 ---GGTLPWGRVGRHSILFNVTFGKELRLRLPNRRL---VVP---GSSVEMQEDERE 142

```

```

QY 127 -----RSHCFYRGTVDSPPSLAVFDLCGLDGFFAVKAHARYTLKPLLRGMAEKEGR 180
Db 143 LFRQPLROECVYTGVTGMPGAANAISNCDGLIRLTDSDFTLEPLERQOEAEASGR 202
QY 181 VYDGSARILHYVTRGFSFELPPRASCETPASTPEAHEHAHNSPGRALASOLL 239
Db 203 TH-----VYRREAAVQOEWAEPDG-----DLHNEAFGIGDLPNLLGLVGDDLG 245
QY 240 DOSALSPAGSGSPQWMMRRRSISRAROVELLVADASMARLYGR-GLQHYLLTLASIA 298
Db 246 DTE-----KRRRAKPGSYSLVLLVVDOSVAFHKEKEVQNVNLTLMIV 291
QY 299 NRYLHASIEHNIIRLAVVAVVVLGDKD-SLEVSNAATTLKFNCKQHONQDGDHDE 357
Db 292 DEYHDESLGIVINIALVRLNWGRQSLIERKPNPSRLSEYCRMASHSQOROPSHAE 351
QY 358 HYDAAILFTREDLQGHSCDTLGMADVGTICSPERSCAVIEDDGLHAAPYVAHEIGHLG 417
Db 352 HHDVAVFLTRDP-----GPGYAPVTGMCHPLRSCALNHEDGFSAFVIAHETGYLG 405
QY 418 LSHD--DSKFC--EETFGSTEDKRLMSILTSIDASKPWSKCTSAITTEFLDDGHCNCLL 474
Db 406 MEHDGNGCCADETSLGS---VMAPLVQAAPFRHFWMSRCKLELSRYLPS--YDCLLD 458
QY 475 LP 476
Db 459 DP 460

```

## RESULT 9

```

O8RE55 PRELIMINARY; PRT: 1223 AA.
AC O8RE55;

```

```

DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Metalloprotease-disintegrin protease.
GN ADAMTS14.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP TISSUE=FETAL LUNG;
RX MEDLINE=21856482; PubMed=11867212;
RA Cal S., Odaya A.J., Llamazares M., Garabaya C., Quesada V.,
  Lopez-Otin C.;
  "Cloning, expression analysis, and structural characterization of
  seven novel human ADAMTSs: a family of metalloproteinases with
  disintegrin and thrombospondin-1 domains.";
  EMBL; AJ345098; CAC87943.1;
  Gene 283:49-62(2002).
DR Integrin; Protease.
KW INTEGRIN.
SQ SEQUENCE 1223 AA; 133925 MW; D585B6593977ED15 CRC64;

```

```

Query Match 16.0%; Score 418; DB 4; Length 1223;
Best Local Similarity 27.6%; Pred. No. 3.9e-25;
Matches 133; Conservative 67; Mismatches 176; Indels 106; Gaps 20;

```

```

QY 21 GPAATPAQDK-AGOPTAAAAQPRRQGEVQEAPEPGHPHPLAQRSSKGLVONIDQ 79
Db 59 GPAASAGSMVVDTPTL-----PRHSHLRV---ARSLHP-----92
QY 80 LYSGG-----GKVG-----YLYVAGRRFLDLERDGSVGIACFVAGGTSAPMWH-----126
Db 93 ---GGTLPWGRVGRHSILFNVTFGKELRLRLPNRRL---VVP---GSSVEMQEDERE 142
QY 127 -----RSHCFYRGTVDSPPSLAVFDLCGLDGFFAVKAHARYTLKPLLRGMAEKEGR 180
Db 143 LFRQPLROECVYTGVTGMPGAANAISNCDGLIRLTDSDFTLEPLERQOEAEASGR 202

```

OY 161 VYGDSSARILHYTREGSFELPRASCETPASTPEAHEHAPANSN-PSGRAALASQL 239  
DB 203 TH-----VYREAAVOEWAEPDG-----DLHNEFGLDLPNLLGLDGOIG 245  
OY 240 DOSALPAGSGPOTWRRRRRSISRAROVELLVADASMARLYGR-GLOHYLTLASIA 298  
DB 246 DTE-----KKRIAKPGSISIEVLLVDDSVYAFHKEKHONVYLTLMNIV 291  
OY 299 NRYLSHASIENHILRAYVKKVVLGDKDK-SLEVSNAATTLKFNCKWOHONLGDDEE 357  
DB 292 DEIYHDESLGVIHINIALVRLIMVYRQSLTERCNPSRLQVCWMAHQRODPRAE 351  
OY 358 HYDAAILFTREDLCGHSCDTLGMADVGTICSPERSCAVIEDDGLAAFTVAHEIGHLG 417  
DB 352 HHDDVFLTRDGF-----GPGYAPVTGMCPLRLSCALNHEDGSSAFVIAHETGVLG 405  
OY 418 LSHD-DSKFC--EETFGSTEDRLMSILTSIDASKPMSKTSATITFEILDGHCNCLD 474  
DB 406 MHHDDGAGCADADETLGS-----VVAPLYQAALFHFHMSRCKLELSKPLPS--YDCLLD 458  
OY 475 LP 476  
DB 459 DP 460

## RESULT 10

OQTEY8 PRELIMINARY: PRT: 1159 AA.

AC OQTEY8: 01-JUN-2002 (TREMBLrel. 21, Created)  
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE ADAMTS14.  
GN ADAMTS14.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
CX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21839041; PubMed=11741898;  
RA Collge A., Vandenbergh I., Thiry M., Lambert C.A., Van Beeumen J.,  
RA Li S.W., Prockop D.J., Lapierre C.M., Nusgens B.V.;  
RT "Cloning and Characterization of ADAMTS-14, a Novel ADAMTS Displaying  
RT High Homology with ADAMTS-2 and ADAMTS-9".  
RL J. Biol. Chem. 277:5756-5766(2002).  
DR EMBL; AF366351; AAL79814.1;  
SQ SEQUENCE 1159 AA; 127336 MW; ASB130149BF7FF34 CRC64;

Query Match 15.7%; Score 410.5; DB 4; Length 1159;  
Best Local Similarity 28.1%; Pred. No. 1.5e-24;  
Matches 125; Conservative 68; Mismatches 173; Indels 79; Gaps 17;

OY 57 PGCHPPLAQRRSKGLVONIDOLYSG---GKVG-----YLVYAGGRFLDLERDGS 107  
DB 6 PPTLP-----RHSHLKVARSPLHPGGLTLPGRGVRGSHLYNTVYFGKELRLRPNR 59  
OY 108 VGIAQFVAGCGTSAWRH-----RSHCFYRGTVDSGRSLAVFDLGGDGEFA 157  
DB 60 L-----SSVEMQEDFRELFRQPLRQCVYTGVTGMPAAVAISNCDGLCLR 112  
OY 158 VNHAYTLKPLLRGPMABEERGRVYGDGSAIRLHYTREGSFELPRASCETASTPE 217  
DB 113 TDSTDFEFLERGOOEKASGRTH-----VYRRAVOEWAEPDG-----D 155  
OY 218 AHEHAPANSN-PSGRAALASQLDOSALSPAGSGPOTWRRRRRSISRAROVELLVAD 276  
DB 156 LHNFAFGLDLPNLLGLDGLDTE-----KKRIAKPGSISIEVLLVAD 201  
OY 277 ASMARLYGR-GLOHYLTLASIANRLYSHASIEHILRAYVKKVVLGDKDK-SLEVSNA 334  
DB 202 DSVAFHKEKHONVYLTLMNIVDEIYHDESLGVIHINIALVRLIMVYRQSLTERCNP 261

OY 335 ATTLKFNCKWOHONLGDDEEYDAAILFTREDLCGHSCDTLGMADVGTICSPERSC 394  
DB 262 SRSLEQVCWMAHQRODPRAEHHDDVFLTRDGF-----GPGMGYAPVTGMCPLRLSC 318  
OY 395 AVIEDDGLAAFTVAHEIGHLGLSHD-DSKFC--EETFGSTEDRLMSILTSIDASKP 451  
DB 319 ALNHEDGSSAFVIAHETGVLGMEHDGAGCADADETLGS-----VVAPLYQAALFHFH 373  
OY 452 WSKTSATITFEILDGHCNCLDP 476  
DB 374 WSRCKLELSKPLPS--YDCLLDLP 396

## RESULT 11

OQSBX0 PRELIMINARY: PRT: 1688 AA.

AC OQSBX0: 01-JUN-2002 (TREMBLrel. 21, Created)  
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE GHI6393p.  
GN CG6107.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
CX Ephydroidea; Drosophilidae; Drosophila.  
CX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RA Stapleton M., Brockstein P., Hong L., Aghayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farlan D., Frise E.,  
RA George R., Gonzalez M., Guarin H., Kromliller B., Li P., Liao G.,  
RA Miranda A., Mungall C.J., Nuncio J., Pacleb J., Paragas V., Park S.,  
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
RA Celniker S.;  
RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY094716; AAM11059.1;  
SQ SEQUENCE 1688 AA; 189867 MW; 48FB8DD4DE0CA4D2 CRC64;

Query Match 15.2%; Score 396.5; DB 5; Length 1688;  
Best Local Similarity 25.8%; Pred. No. 3.5e-23;  
Matches 138; Conservative 76; Mismatches 208; Indels 113; Gaps 19;

OY 2 LIGNASLLCAFRLLP--AANGRAATP--ADKACQPTAAAOPTRRRROGEEOERAP 57  
DB 170 LIG-----MLIMEHLGLRSANKQEBLPQSTHPLASPPATPRLPRRIIDNTSTDHPP 225  
OY 58 PG-----HHP-----LAQRRSKGLVONIDOLYSGGKVG----- 88  
DB 226 DGLDLDDEEHSAFVMPTRKYVYVSLSEADLYESKRNSDINSFLKESASAFAMTGTYRNM 285  
OY 89 -----YLVYAGGRFLDLERDGS-----VGIAQFVAGCGTSAWRH 127  
DB 266 SNEIDPPIQVNLVFGQDLHLVLRQDASFVNNHSMTHIRILKEEENRGPTEAEAGR 345  
OY 128 S-HCFYRGTVDSGRSLAVFDLGGDGEFAVKNARYTLKPLLRGPMABEERGRVYGDGS 186  
DB 346 HLCGFYSGVEDDPHSMVSISLCGTMGTGYIKTSFALLIQPNRTS-SDEVLRHVRFSQ 404  
OY 187 ARLHYTREGSFELPRASCETPASTPEAHEHAPANSNPSGRAALASQLDOSALSP 246  
DB 405 RNARAHSK-----FEL-----GLDDFMSKL--EVOVE 431  
OY 247 AGSGPOTWRRRRRSISRAROV-----ELLVADASMARLYGRGLOHYLTLASIANRLYS 303  
DB 432 ECKSKSRKLNRKRYADVNDQVYTLLEVLIANDNSMKPFGHGDLPYIILMSIYSIFA 491  
OY 304 HASIENHILRAYVKKVVLGDKDKSLEVSNAATTLKFNCKWOHONLGDDEEYDAI 363  
DB 492 DASIONSTRILVRLISLPINIDOTHSSNE--MLKHFCQF--INOSGYER--DTAM 541  
OY 364 LFTREDLCGH---HSCDTLGMADVGTICSPERSCAVIEDDGLAAFTVAHEIGHLGLSH 420

```

Db      542 LITRPGICGVYKICHHMLGLAELGTVCSSSCSIQVDTGIPFAFMALHELILNMNH 600
Qy      421 DSKRCCEFFSGTDEKCR---LMSLITSLDASKPWSKCSATITEFDLGDHCNCL 472
Db      601 DDDKCMPTVTRQNNKVKVLIHMSVWGIHMHPSWKSCKSRHPSLEKTDKSC 655

RESULT 12
O8TE60  PRELIMINARY: PRT: 1081 AA.
AC      O8TE60:
DT      01-JUN-2002 (TREMBLrel. 21, Created)
DT      01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DE      01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE      ADAMTS18 protein.
CN      ADAMTS18.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=21856482; PubMed=11867212;
RA      Cal S., Odaya A.J., Llamazares M., Garabaya C., Quesada Y.,
RA      Lopez-Otin C.,
RT      "Cloning, expression analysis, and structural characterization of
RT      seven novel human ADAMTS, a family of metalloproteinases with
RL      distinct and chondrodomain-1 domains."
RL      EMBL; AJ311903; CAC83612.1;
SQ      SEQUENCE 1081 AA; 119655 MW; 0438BF645676461E CRC64;

Query Match      15.0%; Score 392; DB 4; Length 1081;
Best Local Similarity 27.2%; Pred. No. 4.4e-23;
Matches 149; Conservative 68; Mismatches 224; Indels 106; Gaps 24;

Qy      6 ASLLCAEFLPLAAVCP---AATPAODKAGQ---PPTAAAPRRROGEEVQ--- 53
Db      4 ALLLCAAF---PAAGSGPRGLAGLGKVAKALDCLCCCAVAAALASDSSGASGINDY 61
Qy      54 -----RAEPGH--PHPLAQRRRSGKLVONIDOLYSGGKGVLYVAGRRFLDLERD 105
Db      62 VEVTVEVDASQSYSHDLHNGRKRSQON-----ARSSLHYRSARGQELHLEKPS 115
Qy      106 GSVG---LGFPRAGGCTAPWRHSHCHYRGTVDSPPSLAVFDLCGLDGFPAVKAR 162
Db      116 ALLSHFTIVQVLGKDGASETQKPEVQOCFYQGFIRNDSSSVAVSTCAGLSGLIRTRNE 175
Qy      163 YTLKPL-----LNGPMAEEKGRVYG--DGSARILHYVTRGFSPE 201
Db      176 FLISPLPOLLAQEHNVSSPAGHHPLVYKRTAEKQKRGYRGSQR-----NYRQYSPS 230
Qy      202 ALPPRA-SCETPASTPEAHENAPHSNPSGRAALASQLLDGALSFA-----GG 249
Db      231 HTPHQSOSRET-----EYHHRILQKHFCGR-----KKAAPKPTEDYLFREDEGS 278
Qy      250 SCQPQMMRRRRRSISRAR---VELLVADASMARLYSGN--LQHYLLTLASTANRLYSNA 305
Db      279 SG-----RPRRSAGKSQGLAVETLVADKMKVEKHGKGVNTYTLVYMKVSG--LEKCG 331
Qy      306 STENHRLAVVVVVLGDKDSLEVKNAATLTKNCKQHOHNOIGDDEHYDAALF 365
Db      332 TIGSDINVVVYVLLLEDPGGLLNHNHADQSLNFCQMSA--LIGKNGKH--DAIILL 388
Qy      366 TREDDLCG--HNSCDPLGMAADVGTICSPERSCAVIEDGILHAAPVAHEIGHLGLSHD-D 422
Db      389 TGFDDISWKNNECDLTGFPATSGMCKSKYSCITNEDTGLGFLTAHESGHNFGMIHDE 448
Qy      423 SKFCEETFGSTEDKRLMSLITSLDASKPWSKCSATITEFDLGDHCNCLLPRQILG 482
Db      449 GNPCKRAEGRN-----TMSPLTLGNNGVFSWSSCSROYLTKFLSTPOAGCLVDPKQ---A 500

```

```

Qy      483 GDYKDD 489
Db      501 GQYKYP 507

RESULT 13
O9VF61  PRELIMINARY: PRT: 1229 AA.
ID      O9VF61
AC      O9VF61:
DT      01-MAY-2000 (TREMBLrel. 13, Created)
DT      01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE      01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE      CG6107 protein.
CN      CG6107.
OS      Drosophila melanogaster (Fruit fly).
OC      Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC      Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC      Ephydroidea; Drosophilidae; Drosophila.
OX      NCBI_TaxID=7227;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-BERKELEY.
RX      MEDLINE=20196006; PubMed=10731132;
RA      Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA      Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA      George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA      Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA      Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA      Wan K.H., Doyle A., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA      Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA      Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA      Borkov D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
RA      Butris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA      Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA      de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA      Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA      Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA      Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA      Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA      Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA      Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA      Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA      Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA      Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA      Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA      Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA      Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA      Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA      Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA      Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA      Shue B.C., Siden-Kimmos I., Simpson M., Skupski M.P., Smith T.,
RA      Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA      Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA      Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA      Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA      Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA      Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA      Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT      "The genome sequence of Drosophila melanogaster."
RL      Science 287:2185-2195(2000).
DR      EMBL; AE003709; AAF55199.1;
DR      Flybase; FBgn0038340; CG6107.
DR      InterPro; IPR002870; Rep_M12B_propep.
DR      InterPro; IPR001590; RepPolysin.
DR      InterPro; IPR000884; TSP1.
DR      Pfam; PF01421; RepPolysin.
DR      Pfam; PF01562; Rep_M12B_propep.
DR      Pfam; PF00090; TSP1.
DR      SMART; SM00209; TSP1; 3.
DR      PROSITE; PSS0215; ADAM_MEPRO; 2.
DR      PROSITE; PSS0092; TSP1; 2.

```

DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
KW Hydrolase; Metalloprotease; zinc.  
SQ SEQUENCE 1229 AA; 138136 MW; C1293180AB19B700 CRC64;

Query Match	14.1%	Score 367.5;	DB 5;	Length 1229;
Best Local Similarity	25.5%	Pred No. 5.2e-21;		
Matches 134;	Conservative 72;	Mismatches 206;		
			Indels 113;	Gaps 19

[illegible]

RESULT	14
095N23	
ID	Q95N23
AC	Q95N23;
DT	01-DEC-2001 (TReMBLrel, 19, Created)
DT	01-DEC-2001 (TReMBLrel, 19, Last sequence update)
DT	01-JUN-2002 (TReMBLrel, 21, Last annotation update)
DE	. Aggracnase-2 (Fragment).
OS	Equis caballus (Horse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
OC	Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX	NCBI_TaxID=9796;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Plannery C.R., Little C.B.;
RT	"Expression and activity of equine aggracnases.";
RL	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF368322; AAK53426.1;"
DR	Interpro: IPR001590; Reprolysin.
DR	InterPro: IPR000884; TSP1.
DR	Pfam: PF00090; TSP_1.1.
DR	PROSITE: PSS0215; ADAM_MEPRO: 1.
DR	PROSITE: PSS0092; TSP1.1.
FT	NON_TER
FT	NON_TER
SO	SEQUENCE
	187 AA; 20477 MW; 1078509537ZBEC CRC64;

Query Match 13.5%; Score 354; DB 6; Length 187;

Best Local Similarity 98.5%; Pred. No. 4.7e-21;  
Matches 67; Conservative 0; Mismatches 1; Indels 0; Gaps 0

<b>OY</b>	<b>415</b>	<b>LIGLSHDSDKSCCEETFGSETEDRKLMSSILTSIDASKPMWCKTSTATITEFLDDGNGNCLLD</b>	<b>474</b>
<b>Dd</b>	<b>1</b>	<b>LGLSHDSKRCENPFGEETEDRKLMSLILTSIDASKPMWCKJSTATITEFLDDGNGNCLLD</b>	<b>60</b>
<b>OY</b>	<b>475</b>	<b>LPRKQLLG</b>	<b>482</b>
<b>Dd</b>	<b>61</b>	<b>LPRKQLLG</b>	<b>68</b>

RESULT 15	080VEF1	PRELIMINARY;	PRT: 899 AA.
ID	080VEF1		
AC	080VEF1:		
DT	01-MAR-2002 (TREMBLrel. 20, Created)		
DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	ADAM 19.		
OS	Coturnix coturnix (common quail).		
OC	Archaeoptera: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:		
OC	Cathartidae: Aves; Neognathae: Galliformes; Phasianidae: Phasianinae:		
CC	Coturnix.		
OX	NCBI_TaxID=9091:		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Lewis S.L., Farlie P.G., Newgreen D.F.:		
RT	"Cloning of ADAM 12/19 in Quail."		
RL	Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL, AF56467; AAL5618.1; ..		
DR	InterPro: IPR001762; DisIntegrin.		
DR	InterPro: IPR000561; EGF-like.		
DR	InterPro: IPR002870; pep_M12B_propep.		
DR	InterPro: IPR001590; Repolysin.		
DR	InterPro: IPR000130; Zn_Mpeptidse.		
DR	Pfam: PF00200; disIntegrin.1.		
DR	Pfam: PF01562; pep_M12B_propep. 1.		
DR	Pfam: PF01421; Repolysin. 1.		
DR	PRINTS: PR00289; DISINTEGRIN.		
DR	PRODOM: PD000664; disIntegrin. 1.		
DR	SMART: SM00050; DISIN. 1.		
DR	PROSITE: PSS0215; ADAM_MEPRO. 1.		
DR	PROSITE: PSS0214; DISINTEGRIN.2; 1.		
DR	PROSITE: PS00166; EGF_2; UNKNOWN.1.		
DR	PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN.1.		
DR	SEQUENCE 899 AA: 97896 MW: EC6CE81C377113B35 CRC64:		

Query Match	13.0%;	Score 339;	DB 13;	Length 899;
Best Local Similarity	26.0%;	Pred. No. 7.1e-19;		
Matches 128;	Conservative 70;	Mismatches 191;	Indels 104;	Gaps 21

[illegible]

```
Db 229 ANYVDKFRSL-----NIRIALVLEIWSNGK-CDCTENPYSTLKSFLAMSSKERY---- 279
OY 355 HEHEDDAILFTREDLCGHSCDTLGMADVCTICSPERSCAY---IEDDGLHAFTVAHE 411
Db 280 YRKH-DNAQLITGVPOG---TTVGLAPVMSMCSDFQSGVNMDSMDAIGVAATIAHE 334
OY 412 IGHILGLSHDSEKFCFEETFGSTEDKRLMSSILTSIDASKPMKCTSATTTTEFLDDGHGNC 471
Db 335 MGHNFQNMNHDSSGCC--TTRAEEGGCI MAS-ATGHPRFVFNQCNRRELEKTLHSGGMC 391
OY 472 LIDLPRKQILGG 483
Db 392 ISNMPDTRMYGG 404
```

Search completed: April 21, 2003, 14:47:27  
Job time : 61.6837 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 14:42:31 ; Search time 29.3419 seconds  
(without alignments)  
1611.966 Million cell updates/sec

Title: US-10-050-200-9

Perfect score: 2614  
Sequence: 1 MLGWSALLLCARFLPLAAV.....LDLPKRIILGGDYKDDDKG 492

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	823	31.5	951	2	T00017
2	658.5	25.2	837	2	T00355
3	493	18.9	2165	2	T21371
4	407	15.6	1205	2	T18517
5	290.5	11.1	903	2	S60257
6	275.5	10.5	826	2	A60385
7	263	10.1	814	2	G02390
8	262.5	10.0	617	2	S48160
9	252	9.6	411	1	HYSNPA
10	247	9.4	478	2	JC4880
11	242	9.3	610	2	JC7530
12	239.5	9.2	481	2	JC4342
13	239	9.1	480	1	A30065
14	238.5	9.1	407	2	S62620
15	237.5	9.1	1444	2	T18856
16	236	9.0	549	2	S48169
17	236	9.0	776	2	S28258
18	233.5	8.9	905	2	S55059
19	232.5	8.9	481	2	S43125
20	232	8.9	609	2	S55270
21	231	8.8	789	2	S28259
22	230.5	8.8	571	2	S24789
23	229	8.8	825	2	S55060
24	227.5	8.7	616	2	A55796
25	219	8.4	478	2	A43396
26	219	8.4	860	2	T16892
27	218.5	8.4	414	1	HYSRAC
28	218.5	8.4	414	2	S41609
29	215	8.2	478	2	J01301

30	214.5	8.2	670	2	I65967	distintegrin-like m
31	209.5	8.0	524	2	S38539	distintegrin-like m
32	208.5	8.0	414	2	S41608	atrolysins B (EC 3.
33	207.5	7.9	419	2	S41607	atrolysins A (EC 3.
34	203.5	7.8	419	2	A59414	metalloproteinase
35	202.5	7.7	660	2	S71949	metalloproteinase
36	201.5	7.7	600	2	I49281	feritin alpha pre
37	194.5	7.4	203	2	A59421	metalloproteinase
38	182	7.0	1042	2	T26644	hypothetical prote
39	174.5	6.7	202	2	JC2550	Ac1 metalloprotein
40	172.5	6.6	200	2	S15111	hemorrhagic factor
41	169	6.5	202	1	HYTVH2	hemorrhagic protei
42	167	6.4	201	1	HYTV2	trimericysin II (E
43	165	6.3	429	2	A42972	coagulation factor
44	164.5	6.3	416	2	A37877	hemorrhagic protei
45	160.5	6.1	202	1	HYSR	rubritysin (EC 3.4

## ALIGNMENTS

```

RESULT 1
T00017
gene ADAMTS-1 protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
C:Accession: T00017
R:Kuno, K.; Lizaas, H.; Ohno, S.; Matsushima, K.
Genomics 46, 466-471, 1997
A:Title: The exon/intron organization and chromosomal mapping of the mouse ADAMTS-1 g
A:Reference number: Z14055; MUID:98110583; PMID:9441751
A:Accession: T00017
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-951 <RUN>
A:Cross-references: EMBL:AB001735; NID:q2809056; PIDN:BAA24501.1; PID:q2809057
A:Experimental source: strain 129SVJ
C:Genetics:
A:Gene: ADAMTS-1
A:Introns: 228/1; 343/3; 388/1; 444/1; 539/3; 602/1; 660/3; 719/2
C:Superfamily: thrombospondin type 1 repeat homology
F:542-598/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match          31.5%; Score 823; DB 2; Length 951;
Best Local Similarity 37.3%; Pred. No. 5.6e-55;
Matches 192; Conservative 60; Mismatches 151; Indels 112; Gaps 12;

OY 1 MLGWSALLLCARFLPLAAVGPATPAODKAGOPPTAAAQAQPRRROGEV---QERAE 56
Db 18 MLL-----LLASITMLLCARCAHGRPTE-----EDELVLPSLERA- 54

OY 57 PRGPHPLAORRRKSGLVQNTDOLYSGGKGYLYVAGRRFLDLERDGSVGAFF--- 113
Db 55 -PGHDSYTTTFLR-----LDLFGQDLHLKLPDPSGFLAFTLLQ 91

OY 114 -VPAGGCTSA---PMRRSHCFRGTVDSPRSIAVFDLGGDGFPAVHARYTLKPL 168
Db 92 TVGSPSEAGHDLPTGDLACHFTSGVINGDPGSAALSLICEGRCAFYLDGEEFFIQP- 150

OY 169 LRGWAEERKGRVYGDGSARILHVYTRGFSFELPPRASCETPASPPEAHEHAPAHSNP 228
Db 151 -----APGVATERLAAVPEESSARPOF---ILRRR 181

OY 229 SGRALASQLLDOSALSPAGSGPOT---WNR-----RRRSISR 265
Db 182 RGSAGCAGVWDETL-PTSDSRPSQNTNRQWPVDRPTPDACKPSGPGSIRKRREVSS 240

OY 266 ARQVELLVADASMARLYRGLOHYLLTASIANRLVSHASIEHRIILAVKVVLLDKD 325
Db 241 PRYVETMLVADQSMADRHGSLKHLTLTSVAARFKRHSIRNSISLVYVKIIVIEEQ 300

OY 326 KSLVEYSKNAATTLNFKQKWHQHNOLDGDDEHYDAAILFTREDLCGHSCDTLGMADVC 385
Db 326 KSLVEYSKNAATTLNFKQKWHQHNOLDGDDEHYDAAILFTREDLCGHSCDTLGMADVC 385

```

Db 301 KGEVTSMAALTLRNFNCNMQHNSPSDRDPEHYDTAILFTRODLGCSHTCDTLGMAVYG 360  
 QY 386 TICSPERSCAVIEDDGLHAFTVAHEIGHLLGSHDDSKFCFCEFTFGSTEDKRLMSILTS 445  
 Db 361 TVCDPERSRSCSVIEDDGLQAFTTAHGLGHVFNPHDARHCASLNVGTDSHLMASMLSS 420  
 QY 446 IDASKPWSKTSATITFEFLDDGHCNCLDLPKQI 480  
 Db 421 LDHSQWSPSCSAVAVYTSFLDNGHCGLMDKPPMPI 455

## RESULT 2

T00355  
 hypothetical protein KIAA0688 - human  
 C:Species: Homo sapiens (man)  
 C>Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 21-Jul-2000  
 C:Accession: T00355  
 R:ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.  
 DNA Res. 5, 169-176, 1998  
 A>Title: Prediction of the coding sequences of unidentified human genes. X. The complete  
 A:Reference number: 214142; MUID:98403880; PMID:9734811  
 A:Accession: T00355  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-837 <15H>  
 A:Cross-References: EMBL:AB014588; NID:93327189; PIDN:BAA31663.1; PID:93327190  
 A:Experimental source: brain  
 C:Genetics:  
 A:Gene: KIAA0688  
 C:Superfamily: thrombospondin type 1 repeat homology  
 F:519-573/Domain: Thrombospondin type 1 repeat homology <THR3>

Query Match 25.2%; Score 658.5; DB 2; Length 837;  
 Best Local Similarity 33.6%; Pred. No. 28/42;  
 Matches 170; Conservative 65; Mismatches 158; Indels 113; Gaps 15;

QY 3 LGASALLCAFRPLAAGPATPAODKAGOPATAAAOAPRRRGEVEQERAEPPGPH 62  
 Db 34 LWWLLLLASLPSARL---ASPL-----PREDEIVPEKLNGSVL-----PGSGT 77  
 QY 63 PLARRRRKGLVONIDQLVSGGKGVLYVAGGRFLDLERDGSVGIGF-----VP 115  
 Db 78 PALLLCR-----LQAFETLLLEEDDSGVQVEGLVYQYIGQAP 116  
 QY 116 AGGTSAPWRHRSCHFYKGTVDGSPRSIAVFDLGG-LDGFPAVKHARYTLKPLRGPA 174  
 Db 117 ELTGAEAP-----GYLTGCHINDPESVASLHMDGALLGLVQYRGAEHLQLEGG--- 168  
 QY 175 EEKRGVYDGSARILHYVTRREGSFELPRASCTP---ASTPEAH---EHAPHSNFS 229  
 Db 169 -----TPRSAGCPGAHILRRKSPA-----S 188  
 QY 230 GRAALASQLLDGASLPAGSGSPQWRRRRRISIRAROVETLLVADAMARLYCGLOH 289  
 Db 189 GCGPFC-----NVKAPLSPSPRP--RAKRPASLSRVEVETLLVADKMAAFHCGALKR 240  
 QY 290 YLLTLASIRLYSHASINENHRLAVVYVLDKXKLEVSNAATTLKNCCKMOHQH 349  
 Db 241 YLLTVAAAKAFKHPISINRPVSLVYTRLVILGSGEGEPQVGSNAOTLRSCAMRGRLN 300  
 QY 350 QUGDDEHYDAAILTFREDLCGHNSCDTLGMAVGTCTSPERSCAVIEDDGLHAFTVA 409  
 Db 301 TPEDSDPDHFDVAIFLTRODLGCVSTDLGMAVGTCDPAPRSACIIVEDDGLQSAFTAA 360  
 QY 410 HEIGHLLGSHDDSKFCFCEFTFGSTEDKRLMSILTSIDASKPWSKTSATITFEFLDDGH 468  
 Db 361 HEIGHVFNMLHDSKPCISLNGPLASTSRHVMVAVMHVPPEPWSPCSAFTFDLNDY 420  
 QY 469 GNCLDLDPKQI-----LGDTYKDD 488  
 Db 421 GHCLLDKPEAPHLPLVTPFGKDYDAD 446

## RESULT 3

T21371  
 hypothetical protein F25H8.3 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
 C:Accession: T21371; T24896  
 R:Gajdasty, S.  
 submitted to the EMBL Data Library, February 1996  
 A:Reference number: 219413  
 A:Accession: T21371  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-2165 <WIL>  
 A:Cross-References: EMBL:269360; PIDN:CAA93287.1; GSPDB:GN00022; CESP:F25H8.3  
 A:Experimental source: clone F25H8  
 R:Gajdasty, S.  
 submitted to the EMBL Data Library, February 1996  
 A:Reference number: 219949  
 A:Accession: T24896  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-2165 <WIL>  
 A:Cross-References: EMBL:269361; PIDN:CAA93288.1; GSPDB:GN00022; CESP:F25H8.3  
 A:Experimental source: clone T13H10  
 C:Genetics:  
 A:Gene: CESP:F25H8.3  
 A:Map position: 4  
 A:Introns: 31/1; 52/1; 135/2; 193/3; 216/1; 266/1; 495/2; 547/3; 584/3; 634/2; 744/1;

Query Match 18.9%; Score 493; DB 2; Length 2165;  
 Best Local Similarity 32.1%; Pred. No. 3.3e-29;  
 Matches 126; Conservative 66; Mismatches 124; Indels 76; Gaps 15;

QY 118 GGTSAAPWRHRSCHFYKGTVDGSPRSIAVFDLC---GGLDGFPAVKHARYTLKPLRGPA 174  
 Db 160 GGNS---RTDPPCIYRAHKGVNH--SYNLCDSEDEGLGMLALPSGTHFVPTISGNGT 215  
 QY 175 EEKRGVYDGSARILHYVTRREGSFELPRASCTPSTPEAHEHAPHSNPGRAAL 234  
 Db 216 EH-----DGASRHQHLVLRK---FDPH-----HFKSPDHLN----- 243  
 QY 235 ASQLLDGASLPAGSGSPQW---W-----RRRRSIRAROVETLLVADAMARLY 283  
 Db 244 -STVNETETTYA-----TWDDQMEDVIERKARSRRANSMDHYEVLVAVDTMYEYH 296  
 QY 284 GGLQHYVLLTLASIRLYSHASINENHRLAVVYVLDKXKLEVSNAATTLKNCCK 343  
 Db 297 GRSLEDYVLTSTVASYAIRHOSLRASINVVYKLIYLTENAGRITONAOOTLQDFCR 356  
 QY 344 WOHQNLQGDDEHYDAAILTFREDLC-GHNSCDTLGMAVGTCTSPERSCAVIEDDGL 402  
 Db 357 WOQVYNDPDDSSVQHNDVAIILTRKDICRSQCKDPLGLAELGTMCMDQSCAILEDNGL 416  
 QY 403 HAFTVAHEIGHLLGSHDDSKFCFCE-----TFGSTE-DK-----RLMSILTSI 446  
 Db 417 SAFTIAHEIGHVFSIPHDERRKSTYMPVNVKVKFQSTFKDFOFNNHIMAPTLLEYN 476  
 QY 447 DASKPWSKTSATITFEFLDDGHC--NCLDLP 476  
 Db 477 THPWSKSPCSAGMLERFLFNNGOTQCLFDP 508

## RESULT 4

T18517  
 procollagen N-endopeptidase (EC 3.4.24.14) I - bovine  
 M:Alternate names: procollagen N-proteinase  
 C:Species: Bos primigenius taurus (cattle)  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T18517  
 R:Collage, A.; Nussgens, B.V.; Lapierre, C.M.  
 submitted to the EMBL Data Library, February 1996  
 A:Description: Cloning of the cDNA of the bovine procollagen I N-proteinase.  
 A:Reference number: 218941

A:Accession: T18517  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1205 <COL>  
 A:Cross-References: EMBL:X96389; NID:e990769; PID:e228215; PIDN:CAA65253.1  
 A:Experimental source: skin  
 C:Genetics:  
 A:Gene: PC I-NP  
 C:Function:  
 A:Description: catalyzes cleavage of the propeptides of type I and II collagens prior to  
 C:Keywords: hydrolase; metalloproteinase

Query Match 15.6%; Score 407; DB 2; Length 1205;  
 Best local similarity 28.8%; Pred. No. 6, 2e-23;  
 Matches 147; Conservative 71; Mismatches 203; Indels 90; Gaps 22;

```

QY 4 GWASLLC-----AFRLPLAANGPATPAODKAGOPTAA--AAOPRR--RQGEVOE 53
DB 6 GAAGRLCPALLLLPLLPADARLAAAADPPGPGQGAERILAVPYRTDQGRVLSH 65
QY 54 RAEPGPHPLAQRRRSKGLVONIDQLYSG-----GKRYGYLYVAGRFRLLDERDGS 107
DB 66 VVSATAPAPGVRTRRAAPQOTFGL---SGGSEDDPGRLFYVTVFQDHLRLRPNAR 121
QY 108 VGIAGFVPAAGGTSAPWRHRS-----HCFYRGTVDSFSLAV-FDLGGGLDGEF 156
DB 122 L-----VAPGATVEMGESGATRYEPPLTGCLLYGQVAGLAESSVALSNGDGLAGLI 174
QY 157 AVKHARYTLKPLRGPAEE-EKGRVYGCOSARILHYVREGSFEPALPPRASCETPAST 215
DB 175 RMEEEFPELEPKGLAAKEAEGRVH-----VYVHR-----PTSRPPLGG 217
QY 216 PEANHEAPAHNSPGRAALASQLLDQSALSPAGSGSPQ---TWRRRRRSISRARQVEL 271
DB 218 PQALD-----TGISADSLD--SLSRALGLEERVNSSRRMRRAADDDYNIEV 264
QY 272 LTVADASMARLYG-RGLQHYLLTLASIANRLXSHASIEHHIRLAVKVVVLG-DKDKSLE 329
DB 265 LIGVDSDVVOFQTEHVOYKLLTLMNIVEIYHDESLGHNINVLVRIILLSTGKMSLI 324
QY 330 VSNNAATITLKNCKMOHNOGLDDEEHYDAALFTREDLCGHSCDTLGMADVCTICS 389
DB 325 ETGNPQSLENNWCRAVYLOOKPDTDEHDHDAIFLTRDPF---GPSGMOGYAPVTGMCH 381
QY 390 PERSCAVIDDDGLHAAFYVAHEIGHLGLSHD-DSKFC--EETFGSTEDKRLMSLITSI 446
DB 382 PVRSCITLNEHDGSSAFVAHIEGHVLSGHEHDCGNGRCGDEVRLGS-----IWAPIVQAA 436
QY 447 DASKPWSKCTSATITEFLDDGHC-NCLLDLP 476
DB 437 FHRFHWSRCSQOELSRYL---HSYDCLLRDP 464

```

RESULT 5  
 S60257  
 Meltin alpha - mouse  
 C:Species: Mus musculus (house mouse)  
 C>Date: 10-Apr-1996 #sequence\_revision 19-Apr-1996 #text\_change 20-Jun-2000  
 C:Accession: S60257  
 R:Tagm1-HIromasa, T.; Sato, T.; Kurisaki, T.; Kamijo, K.; Nabeshima, Y.I.; Fujisawa-Sel  
 Nature 377, 652-656, 1995  
 A:Title: A metalloproteinase-disintegrin participating in myoblast fusion.  
 A:Reference number: S60257; MUID:96026308; PMID:7566181  
 A:Accession: S60257  
 A:Status: preliminary; nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-903 <YAG>  
 A:Cross-References: EMBL:D50411; NID:g1054586; PIDN:BA08912.1; PID:g1054587  
 C:Superfamily: mouse meltin alpha; disintegrin homology  
 F:421-503/Domain: disintegrin homology <DIS>  
 F:349/Active site: Glu #status predicted

Query Match 11.1%; Score 290.5; DB 2; Length 903;

Best local similarity 24.8%; Pred. No. 3, 7e-14;  
 Matches 102; Conservative 59; Mismatches 160; Indels 91; Gaps 13;

```

QY 96 RRFLLDLERDGSVGIAGFVPA-----GGTSAPWRHRSCHFRTVDGSPRLAVFDLCG 150
DB 78 RDLISLERNEGLIANGFETHYLDQGTQVLSLRNTHDCYHGHVQGAASVSLSTS 137
QY 151 GDDGFVAVHARTYTLKPLRGPAEEERKRVYGDGSAIRLHYTRGFSFALPPASCE 210
DB 138 DLRLGTLFMRKNTYSLEPM----- 155
QY 211 TPASTPEANHEAPAHNSPGRAALASQ-----LLDOSALSPAGSGSPQTMWRRRRRSISA 266
DB 156 --KNTIDSKVLPAEEMTNIQGLGSGQHNKSNLTMDEVSP--GTSMRARRKRIETLKMT 211
QY 267 ROVELLLVADASMARLYGRLQHYLLTLASIANRLXSHASIEHHIRLAVKVVVLGDKK 326
DB 212 KYVELIVADNHEFOQGDLEKVKORLIEINHYDKFPRPLN-IRIVLVGEVWMDIDK 270
QY 327 SLEVSNAATITLKNCKMO-----HQNOLGDDEHDAALFTREDLCG-HHSC 376
DB 271 -CSISODPPTRLHEFLDMKRIKLLPRKSHDNOL-----ISGVYFGG 311
QY 377 DTLGMADVCTICSPERSCAVI---EDDGLHAFTVAHEIGHLGLSHD-SKFCETFGS 432
DB 312 TTIGMAPIMSMCTAEQSGGVVDHSDSPILGAATLAHEIGHNFGNMHDTLERGCSGRMA 371
QY 433 TEDKRLMSLITSIDASKPWSKCTSATITEFLDDGHCNCLDLP-KQILGG 483
DB 372 EKGCIIMNP-STGFPPMVFSSCSRKDLBASLEKGMCLFNLPEVKAQFGG 422

```

RESULT 6  
 A60385  
 monocytic surface antigen MS2 precursor - mouse  
 C:Species: Mus musculus (house mouse)  
 C>Date: 03-Feb-1993 #sequence\_revision 03-Feb-1993 #text\_change 26-Aug-1999  
 C:Accession: A60385  
 R:Yoshida, S.; Setoguchi, M.; Higuchi, Y.; Akizuki, S.; Yamamoto, S.  
 Int. Immunol. 2, 585-591, 1990  
 A:Title: Molecular cloning of cDNA encoding MS2 antigen, a novel cell surface antigen,  
 A:Reference number: A60385; MUID:91197896; PMID:1982220  
 A:Accession: A60385  
 A:Molecule type: mRNA  
 A:Residues: 1-826 <YOS>  
 A:Cross-References: EMBL:X13335  
 C:Superfamily: mouse meltin alpha; disintegrin homology  
 C:Keywords: glycoprotein; surface antigen; transmembrane protein  
 F:1-14/Domain: signal sequence #status predicted <SIG>  
 F:402-484/Domain: disintegrin homology <DIS>  
 F:659-683/Domain: transmembrane #status predicted <TM>  
 F:330/Active site: Glu #status predicted

Query Match 10.5%; Score 275.5; DB 2; Length 826;  
 Best local similarity 24.2%; Pred. No. 4, 7e-13;  
 Matches 118; Conservative 62; Mismatches 174; Indels 133; Gaps 20;

```

QY 34 PPTAAAAOPRRRQGEVOERAEPRPHPLAORRSKGLVONIDQLYSGGKVGVLVYA 93
DB 13 PAAVAPPPPLPHVQYEVV-----WPRRLAASRRRALPSHWGQYPE---SLSYALGT 61
QY 94 GGRFFLLDERD-----GSVGIAGFVPAAGG-TSAPWRHSHCFRTVDGSPRLAVFL 148
DB 62 SGHVFLLLRKNLDLGGSTYETYSANQSEVTEOLOEDDHCLYQGHVGEBSAASIST 121
QY 149 CGGLDGFPAV-----KHARYTLKPLRGPAEEERKRVYGDGSAIRLHY 192
DB 122 CAGLRGEFFRVGTVHLIERPLADEQGHMAYQAKHLQKAGTCGVADTNLDGPRALIEI 181
QY 193 YTRGGSFEPALPPRASCETPASTPEANHEAPAHNSPGRAALASQLLDQSALSPAGGSP 252
DB 182 YRAQ-----P 186

```

```

OY 253 OTMMRRRRRRSTSRARQVELLVADASAAIRYUG---RGJONHLLTILASTAIRVSHSXIEMH 310
Db 187 RNMILPR-----ETRIYVELLVADSOEFOGLSGREAVRQVLEVNHHDKLQJOLESF--- 238
OY 311 IRLAVKVVULG----DKDXSLEYSKNAATLTNFKCKMOHONOLCDDHEENYDAALIFT 366
Db 239 -----RYVLEIGETIMNKK-RFYSRVANTLTNPLSMREQNPO----GQHPHDVQILIT 287
OY 367 REDLGHHSODTLGMAVUGTIGSPERSCAITEP---DGLHAATVAHNEGILLGLSHND- 422
Db 288 GVDITG----STVGLAVYSALCS-RHSAGVANOHSKNNSIGVSTMKHEHNLHNGMSHDD 342
OY 423 --SKFCETEGSTEDKRLMSILT-SIDASKP--WSKTSATITTEFLDGHGNCILDIPLR 477
Db 343 IPGCGYC-----PEPRREGGCGCIMTESIGSKFPRIIFRCSKIDLESIVTMRQGLGNLPVD 396
OY 478 -KOLLG 483
Db 397 VNRFEVGG 403

```

## RESULT 7

disintegrin-like metalloproteinase MDC15 (EC 3.4.24.-) - human  
 C:Species: Homo sapiens (man)  
 C:Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 31-Dec-2000  
 C:Accession: G02350; PC4263  
 C:Herren, B.; Ralnes, E.W.; Ross, R.  
 submitted to the EMBL Data Library, January 1996  
 A:Reference number: H01157  
 A:Accession: G02350  
 A:Status: preliminary; translated from GB/EMBL/DBD  
 A:Molecule type: mRNA  
 A:Residues: 1-814 <EMBL>  
 A:Cross-references: EMBL:U46005; NID:G1335871; PIDN:AA05112.1; PID:G1335872  
 R:Miche, N.; Edwards, T.; Dallis, D.J.; Houghton, A.; Stringer, B.; Graham, R.; Russell  
 Biochem. Biophys. Res. Commun. 230, 335-339, 1997  
 A:Title: Expression of members of a novel membrane linked metalloproteinase family (ADAM  
 A:Reference number: PC4263; MUID:9716897J; PMID:9016778  
 A:Accession: PC4263  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-461 <MCK>  
 A:Experimental source: articular chondrocyte  
 C:Comment: This protein is a membrane bound protein and involved in cell/cell and cell/?  
 C:Superfamily: mouse metlatin alpha; disintegrin homology  
 C:Keywords: hydrolase; metalloproteinase; zinc  
 F:420-503/Domain: disintegrin homology <DIS>  
 F:348,352,358/Binding site: zinc (His) #status predicted  
 F:349/Active site: Glu #status predicted

Query Match	10.1%	Score 263	DB 2	length 814
Best Local Similarity	27.8%	Pred. No.	4.2e-12	
Matches 113	Conservative 51	Mismatches 136	Indels 86	Gaps 18

```

Oy 97 RFLDLEND-----GSVGIACGVPVAGCGSASMRHNSHCFTGTVGSPRS 142
Db 70 RIKELDODSHILELQRLVPGRPRLVWYQPDOTRVYSECHTLENCYQGRVAGTGS 129
Oy 143 LAVPDLCCGDLGFFAVKHAR-YTLKPLRLGFWAEENKGVYDGSARILLHYVTRGFSFE 201
Db 130 WVSICTSGLGLVLPERSYTLT---QGP-----GDLOG----- 162
Oy 202 ALPRASETPASTPEAHENHAPHSNP--SGRAALASQLDLSALSPPAGSGPQTWMRRRR 260
Db 163 --PPLISITQDL-----HLPGHTCALSMRESVNTQPPPHPLG-----QRHRRRR 206
Oy 261 RSISAROVLELVADASMARLYGRLQHYL---LTLASINRLYSHASINHIIRLAVVK 317
Db 207 DVVETKTELVELVIAVDHSEAGKY-RDFQHLLNRLLEVALLLDTPFRPL---NVRVALVG 261
Oy 318 VVVVGLDCKDLSLEVSKNNAITLKNECKWOHONQJLDDHNEHYDAIILFTREDLGGHNSCD 377

```

[illegible]

## RESULT & DISCUSSION

metalloproteinase (EC 3.4.24.-) H-I precursor - carpet viper  
N:Contains: disintegrin  
C:Species: Echis pyramidalis leaky  
C:Date: 19-Mar-1997 #sequence\_revision 13-Mar-1997 #text\_change 21-Jul-2000  
C:Accession: S48160  
R:Palme, M.J.I.; Moura-da-Silva, A.M.; Theakston, R.D.G.; Crampson, J.M.  
Eur. J. Biochem. 224: 483-488, 1994  
A:Title: Cloning of metalloproteinase genes in the carpet viper (Echis pyramidalis leaky)  
A:Reference number: S48160; MUID:95010023; PMID:7925363

A: Molecule type: mRNA  
A: Residues: 1-617 <PAI>  
A: Cross-references: GB:K78970; NID:g763092; PIDN:CA55565.1; PID:g763093  
C: Superfamily: mouse metalloprotease; disintegrin homology  
C: Keywords: hydrolase; metalloproteinase; venom  
F: 1-18/Domain: signal sequence #status predicted <SIG>  
F: 15-192/Domain: propeptide #status predicted <PRO>  
F: 193-617/Product: metalloproteinase H-I #status predicted <MAT>  
F: 403-485/Domain: disintegrin homology #status predicted <DIS>  
F: 337/Active site: Glu #status predicted

Query Match	10.0%	Score 262.5;	DB 2;	Length 617;
Best Local Similarly	24.3%;	Pred. No. 3.2e-12;		
Matches 105; Conservative	68;	Mismatches 176;	Indels 83;	Gaps 17

62 HPLAQRRSKGLVQNIQILYSGGKVGILVYAGGRFLDLDERDSV-----GIAGFVPA 116

Db 35 YPQKNA TLTKEA IQP EQKY E--DTMR YE K VNGEPVLHLLEK NKG L FSEDSY SETHY SPD 92

[illegible][illegible]

OV 237 OLDOSALPAGSGPOTWBRBRBSISRABOVETLTADASMRLYGR--GIQHYLT 293

Db 187 QLVATS-----EQORSYK--KYIEFVVADYIMYRKYNNDSTAVRRIYE 229

QY 294 LASIANRLYSHASIEHIRLAVKVVLGDKDKSLEVSNAATTLKNFCKWQHQNQLGD 353

Db 230 I V N I L M V - - - - I V F N I H A L T H I E I W S T R D Q - I T V Q S A D V T L D L F G D W R A K N L L T R K 284

334 DHEENIDANLEIKEDCGHNSCDLGMAUVG1CSPEKSCAVIED--DCLNAAFIVAH 410

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

[illegible]

OV 467 GHGNCIT.DI.PRK 478

Db 387 DQSKCIDNKPLK 398

## HYSNFA

HYSNEFA

















Query Match 29.4%; Score 769.5; DB 9; Length 823;  
 Best Local Similarity 42.8%; Pred. No. 1.6e-55;  
 Matches 180; Conservative 49; Mismatches 133; Indels 59; Gaps 12;

QY 82 SGGGKGVLYVAGRRFLDLLENDGSGVTAAGF-----VPAGG--GTSAPWRHRSCTFR 133  
 DB 48 SGGGGLFDTTAQEDFYHLTPDAQFLAPAFSTEHLGVPLDGLTGSSDLR---RCFYS 104  
 QY 134 GTVDGSPRLAVDLGCGLDGFFAVHARHTLRLRGPAEERKGRVYGDGSA RLTHY 193  
 DB 105 GGVNAEPDSEAAVSLCGGLGAGFYGAELVSP---PRASAPAAQRNSQGA---HLL 157  
 QY 194 TREGSEFALPPRASCETPASTPEAHEHAPAHNSPGRAALAS---QLDOSALSP--- 246  
 DB 158 ORRG-----VPCGPS-----GDPTSRGVASGMPAL--FALDPKP 193  
 QY 247 -AGSGPOTWRRRRS-----ISRAQVELLVADASMARLYSGLOHYLLTLASIN 299  
 DB 194 RRAFGES---RSRRSGRAKRFVSIPIRYETLVADSVKFGADLEHYLLTLATAA 250  
 QY 300 RLYSHASIEHRLAVKVVVLGDKDKSLEVSNAATTLKNECKWOHNOGLDDEHY 359  
 DB 251 RLYRHSLINPIVYVYKLLNDRSGPKYGNALTLRNFCAKMKLKYSDKHPEY 310  
 QY 360 DAALFTREDLCGHSCDPLGMAADVCTICSPERSCAVIEDDGLAAFTVAHEIGHLGS 419  
 DB 311 DVALFTRODLCATCTDPLGMADVCTMCDPKRSCVIEDDGLPSAFTTAHEIGHVFNMP 370  
 QY 420 HDSKCEETFGSTEDKRLMSSILTSIDASKPMKCTSAITFEFLDGGNCLLDPRKO 479  
 DB 371 HDVAKCEEFGLKRLANHMSPILQIDRANPWSACSAAILTFDLDGHDCLLDPSKP 430  
 QY 480 I 480  
 DB 431 I 431

RESULT 7  
 US-09-918-171A-9  
 ; Sequence 9, Application US/09918171A  
 ; Patent No. US20020110894A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Apte, Suneel  
 ; APPLICANT: Hurskainen, Taina L.  
 ; APPLICANT: Hirahata, Satoshi  
 ; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases  
 ; FILE REFERENCE: 26473/04193  
 ; CURRENT APPLICATION NUMBER: US/09/918,171A  
 ; CURRENT FILING DATE: 2001-07-30  
 ; PRIOR APPLICATION NUMBER: 09/369,364  
 ; PRIOR FILING DATE: 1999-08-06  
 ; NUMBER OF SEQ ID NOS: 31  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 9  
 ; LENGTH: 905  
 ; TYPE: PRT  
 ; ORGANISM: Mus  
 US-09-918-171A-9

Query Match 28.0%; Score 732.5; DB 10; Length 905;  
 Best Local Similarity 33.5%; Pred. No. 2.1e-52;  
 Matches 177; Conservative 61; Mismatches 185; Indels 105; Gaps 11;

QY 4 GNASLLCAFRPLPLAVGPAATPAODKAGOPTAAQAOPRRRQGEVEGERAEP---PGH 60  
 DB 9 GMPPLLLLLQLP-----PPPLVCGAAGAGPCTGQASELVPTPLPS 51  
 QY 61 PHPLAQRARRSKGLVONIDLYSGGKGVLYVAGRRFLDLLENDGVS-----GIAG 112  
 DB 52 ASBLA-----FHLAIFGGGFVLRALAPDASFLAPFKIERLGG 88  
 QY 113 FVPAGGCTAPWRHRSCTYRGTVDSPPSLAVFDLCGGLDGFFAVHARHTLTKPLLRP 172

DB 89 SSAAAGBERG---LRCFSGTVNGERESLAAMSCVAMGSSFLLAGEEFTIOPGACD 144  
 QY 173 -----WAEERKGRVYGDGSA RLTHYTRREGSFALPPRASCETPASTPEAHEHA 222  
 DB 145 SLDDPRLORMGCGRRREDPLAAAEV-----PLDGLWEVEMKNGGQERS 193  
 QY 223 P-AHNSPGRALASQLLDQSAISPAGSGPOTWRRRRRSISRAQVELLVADASMAR 281  
 DB 194 DNEEDKKQKKEGLKTEDESRKVPPEFGSK-----TRSRFPSEAEVETLLVADASMA 248  
 QY 282 LVGRGLQHLTLTLASTANLVSASIEHNI RLAVKVVVLGDKDKSLEVSNAATTLKNE 341  
 DB 249 FCTDQDNLHILTVSMAARIYKHPISIRNSVNLVYKVLIVEREMKPEVSDNGGLTRNF 308  
 QY 342 CKWOHNOGLDDEHYAAIIFTRDLCGH-HSCDTLGMADVCTICSPERSCAVIEDD 400  
 DB 309 CSNORFKNPDRPHRYTALITFTQNCCKGECDDTLGMADVCTICPPDKSCSYINKE 368  
 QY 401 GLHAFTVAHEIGHLGLSHDSKFCETFGSTEDKRLMSSILTSIDASKPMKCTSATI 460  
 DB 369 GLQAAVTLAHELGHVLSMPHDDSKPCVRLFGPKYHMAAPFIIHVNKTLPWSPCSAVYL 428  
 QY 461 TERLDGHCNCLLDLPR-----KQILGDYK 486  
 DB 429 TELDDGHDCLLDAPTSVLPPTGLGHSYLYELDQCKQIFGPDR 476

RESULT 8  
 US-10-097-597-12  
 ; Sequence 12, Application US/10097597  
 ; Publication No. US2003002232A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hirose, Kunitaka  
 ; APPLICANT: Inoguchi, Eiji  
 ; APPLICANT: Hakozaaki, Michinori  
 ; APPLICANT: Ishioaka, Keiko  
 ; APPLICANT: Ishida, Yukako  
 ; APPLICANT: Matsushima, Kouji  
 ; APPLICANT: Kuno, Kouji  
 ; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same,  
 ; TITLE OF INVENTION: Pharmaceutical  
 ; TITLE OF INVENTION: Composition and method of immunologically analyzing human ADA  
 ; FILE REFERENCE: 057092  
 ; CURRENT APPLICATION NUMBER: US/10/097,597  
 ; CURRENT FILING DATE: 2002-03-15  
 ; PRIOR APPLICATION NUMBER: 09/445,023  
 ; PRIOR FILING DATE: 1999-12-03  
 ; PRIOR APPLICATION NUMBER: JP 9-160422  
 ; PRIOR FILING DATE: 1997-06-03  
 ; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 12  
 ; LENGTH: 727  
 ; TYPE: PRT  
 ; ORGANISM: Mus sp.  
 US-10-097-597-12

Query Match 27.7%; Score 723; DB 9; Length 727;  
 Best Local Similarity 57.6%; Pred. No. 9.8e-52;  
 Matches 133; Conservative 35; Mismatches 61; Indels 2; Gaps 1;

QY 250 SGPOTWRRRRRSISRAQVELLVADASMARLYSGLOHYLLTLASINRLYSHASIE 309  
 DB 3 SGPGS--IRKKRRVSSPRVETMLVADQSMADPHSGGLKHYLLTLESVAARFYKHPISIN 60  
 QY 310 HIRLAVKVVVLGDKDKSLEVSNAATTLKNECKWOHNOGLDDEHYDAALIFTRD 369  
 DB 61 SISLVYKILVYEEQKGEVTSNALTLRNFCSMKQHNPSPSDRPDEYDTAALIFTRD 120  
 QY 370 LCGHNSCDPLGMADVCTICSPERSCAVIEDDGLAAFTVAHEIGHLGLSHDSKRCET 429  
 DB 121 LCGSHCTDPLGMADVCTCDPSRSCVIEDDGLQAAFTTAHEIGHVFNMPHDAKHCASL 180



Db 129 TLGMADVTVCDDPSRSCSYVIDDDQLQAFTTAHELGAVFENKPHDDAQCQASLNGVNDSH 188

Qy 438 LMSILITSIDASKPMSCKTSATITFEPLDDGHCNCLLDPKQI-IGGD 484

Db 189 MMASRLSLNDHSQPMSPCSAAYMTISFLDNGHGECILMKPQNPPIPLPGD 236

```

RESULT 12
US-10-097-580-1
: Sequence 1, Application US/10097580
: Publication No. US20030032168A1
: GENERAL INFORMATION:
: APPLICANT: Hirose, Kunitaka
: APPLICANT: Inoguchi, Eiji
: APPLICANT: Hakezaki, Michinori
: APPLICANT: Ishioke, Keiko
: APPLICANT: Ishida, Yukako
: APPLICANT: Matsushima, Kouji
: APPLICANT: Kuno, Kouji
: TITLE OF INVENTION: Human ADAMS-1 protein, gene encoding the same, pharmaceutical
: TITLE OF INVENTION: composition and method of immunologically analyzing human ADAMS-1
: FILE REFERENCE: 051092
: CURRENT FILING DATE: 2002-03-15
: CURRENT FILING DATE: 2002-03-15
: PRIOR APPLICATION NUMBER: 09/445,023
: PRIOR FILING DATE: 1999-12-03
: PRIOR APPLICATION NUMBER: JP 9-160422
: PRIOR FILING DATE: 1997-06-03
: NUMBER OF SEQ ID NOS: 14
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 1
: LENGTH: 727
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-097-580-1

```

```

RESULT 13
US-09-445-023A-1
; Sequence 1, Application US/09445023A
; Patent No. US20020119167A1
; GENERAL INFORMATION:
; APPLICANT: Hirose, Kunitaka
; APPLICANT: Inoguchi, Eiji
; APPLICANT: Hakozeaki, Michimori
; APPLICANT: Ishioka, Keiko
; APPLICANT: Ishida, Yukako
; APPLICANT: Matsushima, Kouji
; APPLICANT: Kuno, Kouji
; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical
; FILE REFERENCE: Q57092
; TITLE OF INVENTION: composition and method of immunologically analyzing human ADAMTS

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

RESULT 14
US-10-174-590-352
; Sequence 352, Application US/10174590
; Publication No. US20030008352A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C42
; CURRENT APPLICATION NUMBER: US/10/174,590
; CURRENT FILING DATE: 2002-06-18
; Prior application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 352
; LENGTH: 837
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-590-352

Query Match      25.2% Score 657.5; DB 9; Length 837;
Best Local Similarity 33.9% Pred. No. 3.2e-46;
Matches 172; Conservative 63; Mismatches 156; Indels 117; Gaps 15.

      3  LGMASLLCAFLPLAANGPAATPAQDKAGOPTTAAAAQPRRRGGEYOEKAEPPGPH 62
      |  |||  |  |
      34  LVMLLLLLLASLPL-----SARLASPLPREEEIV-----FPE 65
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY      63  PLAQRNRSGLVONIDOLYSGGKGYL--VYAGGRFLIDLRGDSVGIAGF----- 113
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB      66  KLV-----NCSVLPGSGAPARLLCRQAFETLLLEEDSGVGEGLVQYLIQ 114

```

[illegible]

RESULT 15.  
US-10-176-758-352  
; Sequence 352, Application US/10176758  
; Publication No. US20030008353A1

; GENERAL INFORMATION:

```

? APPLICANT: Baker, Kevin P.
? APPLICANT: Chen, Jian
? APPLICANT: Desnoyers, Luc
? APPLICANT: Goddard, Audrey
? APPLICANT: Godowski, Paul J.
? APPLICANT: Gurney, Austin L.
? APPLICANT: Pan, James
? APPLICANT: Smith, Victoria
? APPLICANT: Watanabe, Colin K.
? APPLICANT: Wood, William I.
? APPLICANT: Zhang, Zemin
? TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
? FILE OF INVENTION: ACIDS ENCODING THE SAME
? FILE REFERENCE: P3430R1C104
? CURRENT APPLICATION NUMBER: US/10/176,758
? PRIOR FILING DATE: 2002-06-21
? Prior Application removed - See file Wrapper or Palm
? NUMBER OF SEQ ID NOS: 612
? SEQ ID NO 352
? LENGTH: 837
? TYPE: PRT
? ORGANISM: Homo Sapien
? US-10-176-758-352

```

Query Match	25.2%	Score 657.5	DB 9	Length 837
Best Local Similarity	33.9%	Pred. No. 3.2e-46		
Matches 1/2; Conservative	63	Mismatches 156	Indels 117	Gaps 15

```

QY      3 LGMSLLLCAPRLPLAAGPATGATQXAGRPFAAAAAAORRRKGGEGEORERAPRRNPH 62
      4 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db      34 LVMLLLLLLASTLP-----SARLASLPFREEIV-----PPE 65

QY      63 PLAORRSRSGIVONIDOLYSGGKGVYL--VYAGGRFLDLEDGSGVYAGF----- 113
      66 KL-----KNSVLPGSGAPARLLCRIGANGETLLLEDEDSGVQYEGLVYQIGQ 114

QY      114 VPAGGCTAPMRHNSHCYRGTVDGSPSLAVFDLGG-LDGFPAVKHARYTKPLRLGP 173
      115 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

```

Db 115 APELLGAEPE-----GTYLTGTINGDEPESVASHMDGALLGVLOQYGAELHLOPLEGG- 168
QY 173 WAEBEKGAVYDGSARILIAVITRECGSEFALPPRASCETP--ASTPEAH---EHAAPHSN 227
Db 169 -----TPNSAGCPGAHTILRRKSPFA--- 187
QY 228 PSGRALIASQLLDDOSALSPAGSGPOTMMRRRRSISRARQVELLYVADASMARLYRGU 287
Db 188 -SGGCPMC-----NVKAPLGLSPSPRP--RAKRFPA$LSRVELLYVADKMAAFHAGL 238
QY 288 QHYLLTTLASIANRLYSASHASINENHRLAVKVVVGLDGDKSLSEVSKNAATLKNPCQKHO 347
Db 239 KRYLLTVMAAAAKAFKRPISINP$SLVYTRLYLVLGSEBGPQVGP$AAO$TLR$CANORG 298
QY 348 HNOLDDHEEHYDAALFETREDLGGHNSCDTLGNADVGTICSPERSCAVIEDDGLAAFT 407
Db 299 INTPEDSGPDHFOTAILFETRDLCGVSTCDTLGNADVGTCDPARSCAIVEDDGLQ$AFT 358
QY 408 VAHEHGLILGISHDSCFCEETFE$TEDEKR-LMS$ILIT$IDASK$PM$K$C$ATITTEFLD 466
Db 359 AAHELG$HVE$NNH$JHNSKPC$IL$N$PL$T$SRH$V$A$P$V$A$H$V$D$PE$P$M$S$C$AR$F$IT$DEL$N 418
QY 467 GHGNCILLDLP$R$KQJ-----LGGGYKDD 488
Db 419 GYGHCLLDKPEAP$PL$V$Y$T$PE$G$K$OT$D$AD 446

```

Search completed: April 21, 2003, 14:55:54  
Job time : 24.2907 secs

Job time : 24.2907 secs

THIS PAGE BLANK (USPTO)



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 14:40:26 ; Search time 53.3163 Seconds  
(without alignments)  
1727.484 Million cell updates/sec

Title: US-10-050-200-8  
Perfect score: 2370  
Sequence: 1 MSQTSHPGRGLAGRLMGA.....PEAPLPLPVTGDPKDDDDKG 447

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SP-archaea:\*  
2: SP-bacteria:\*  
3: SP-fungi:\*  
4: SP-human:\*  
5: SP-invertebrate:\*  
6: SP-mammal:\*  
7: SP-mhc:\*  
8: SP-organellar:\*  
9: SP-phage:\*  
10: SP-plant:\*  
11: SP-rodent:\*  
12: SP-virus:\*  
13: SP-vertebrate:\*  
14: SP-unclassified:\*  
15: SP-virus:\*  
16: SP-bacteriap:\*  
17: SP-archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	896.5	37.8	950	4	08TE58	08te58 homo sapien
2	577	24.3	192	6	095N24	095n24 equus caball
3	532	21.6	5	019791	019791 caenorhabdi	
4	457	19.3	1054	5	09W493	09w493 drosophila
5	455	19.2	1095	4	08TE56	08te56 homo sapien
6	418.5	17.7	1081	4	08TE60	08te60 homo sapien
7	394	16.6	1207	4	08TE59	08te59 homo sapien
8	393	16.6	1688	5	08SX80	08sx80 drosophila
9	384	16.2	1072	4	08TE57	08te57 homo sapien
10	382	16.1	117	11	08VHK4	08vbk4 mus musculu
11	353.5	14.9	1229	5	09V61	09v61 drosophila
12	341.5	14.4	1223	4	08WXS8	08wxs8 homo sapien
13	341.5	14.4	1223	4	08TE55	08te55 homo sapien
14	340	14.3	1159	4	08TEY8	08tey8 homo sapien
15	312.5	13.2	1091	5	09W126	09w126 drosophila
16	307.5	13.0	269	6	09GL54	09gl54 oryctolagus

17	304	12.8	1427	4	096137	096137 homo sapien
18	251.5	10.6	790	5	08T458	08t458 drosophila
19	248.5	10.5	407	13	092032	092032 agkistrodon
20	247.5	10.4	409	13	080G89	080g89 bothrops in
21	238.5	10.1	477	13	088SP2	088sp2 bothrops ja
22	236.5	10.0	604	13	09PT48	09pt48 atracaspis
23	234	9.9	505	13	073795	073795 agkistrodon
24	233.5	9.9	482	13	09PVK9	09pvk9 agkistrodon
25	228.5	9.6	480	13	090220	090220 agkistrodon
26	228	9.6	479	13	09PMU0	09pmu0 agkistrodon
27	227.5	9.6	922	13	08UVF2	08uvf2 coturnix co
28	226	9.5	899	13	08UVF1	08uvf1 coturnix co
29	225.5	9.5	476	13	09Y119	09y119 agkistrodon
30	223.5	9.4	483	13	09IAB0	09iab0 agkistrodon
31	222	9.4	610	13	080G88	080g88 bothrops in
32	221	9.3	400	13	09IAY0	09iay0 agkistrodon
33	220.5	9.3	404	13	090392	090392 cirotatus at
34	220	9.3	479	13	09IAY1	09iay1 agkistrodon
35	220	9.3	478	13	09PW78	09pw78 agkistrodon
36	219	9.2	478	13	090W25	090w25 trimeresuru
37	219	9.2	610	13	093523	093523 bothrops ja
38	218.5	9.2	609	13	09W6M5	09w6m5 agkistrodon
39	218	9.2	466	13	09IAX7	09iax7 agkistrodon
40	218	9.2	466	13	09IAX6	09iax6 agkistrodon
41	216.5	9.1	411	13	092031	092031 agkistrodon
42	216	9.1	481	13	057413	057413 trimeresuru
43	215	9.1	400	13	09IAY3	09iay3 agkistrodon
44	213.5	9.0	487	13	092119	092119 trimeresuru
45	213	9.0	612	13	08UVG0	08uvg0 bothrops er

## ALIGNMENTS

RESULT 1  
ID 08TE58 PRELIMINARY: PRT: 950 AA.  
AC 08TE58;  
DT 01-JUN-2002 (TREMBLrel. 21, Created)  
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Metalloprotease disintegrin 15 with thrombospondin domains.  
GN ADAMTS15.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN 11  
RP SEQUENCE FROM N.A.  
RX MEDLINE-21856482; Pubmed-11867212;  
RA Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V.,  
RA Lopez-Otin C.;  
RT Cloning, expression analysis, and structural characterization of  
RT seven novel human ADAMTSs, a family of metalloproteinases with  
RT disintegrin and thrombospondin-1 domains.\*;  
RL Gene 283:49-62(2002).  
DR EMBL: AJ315733; CAC86014.1; -  
KM Integrin; Protease.  
SQ SEQUENCE 950 AA: 103286 MW: 5DPBE18285CCCC3B CRC64:  
Query Match 37.8%; Score 896.5; DB 4; Length 950;  
Best Local Similarity 43.6%; Pred. No. 3.7e-60;  
Matches 190; Conservative 71; Mismatches 136; Indels 39; Gaps 9;  
OY 37 LLLLLASLLPSARLASPLFREIEVPEKLNQSVL-----PGSGTPARLLCRLOAF 88  
DB 1 MLLGLITTLAFAGRTAGSGSEPEREVVPIRDPDINRRYRWGPEPDSGDGQILFIQTAR 60  
OY 89 GETLLLELDSDSGVQVGLTVQYLGAPE--LLGGAEP--CTYLGTITNGDPESVASLHMD 145  
DB 61 QEDFYHLTPDAQFLAPAFSTEHLGVPLQGLTGSSSLRRCFTSGDVNAEPDSFAVSLC 120  
OY 146 GGALLGVLYRGAEHLQPLEGGTPNSA--GPGCAHLRRK---SPASGGPFCNV--- 196

```

Db 121 GG-LRCAFGRGAEYVISPAPMAORNSOAHLLQRRGVGCGSPDTSKCGVANS 179
197 -----KAPLCSPPRRP--RAKRPASISREVEITLVADDKMAAFHAGL 239
Db 180 WNPAILALDPPYKRRAGFESRRSRGAKRFVSIPTVETLVVADSKVFFHGDL 239
240 RYLLTVMAAAAKAFKHSINRPSLVYLRVLVIGSGEGPOVPSAOTLRSCAMORGL 239
Db 240 HYLTLTATARLYRHSIINPIVYVVKLLLRDRSGPKVGNMALTIRNCAMQKXL 239
300 NTBEDSPDFDPAITLFTRODLCVSTCDPLGADVGTVCDDPARSCAIVEDDGLQSAFTA 359
300 NKVSDKRPEDVDAITLFTRODLCGATTCPLGADVGTMCDDPKRSCSVIEDDGLPSAFIT 359
360 AHELGHVFNALHNSKPCISLNGPLSTRHVA PVMAHVDPPEPSPCSARITDFLDNG 419
Db 360 AHELGHVFNPHDNVKKVCEVFGKL-RANIMMPTLIQIDRANPWSACSAITITDFLDG 418
Qy 420 YGHCLDKPEAPLHP 435
Db 419 HGCDLLDPPSKPISLP 434

```

## RESULT 2

```

ID 095N24 PRELIMINARY: PRT: 192 AA.
AC 095N24;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
OS Aggrecaenase-1 (Fragment).
OC Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_Taxid=9796;
RN [1]
RP SEQUENCE FROM N.A.
RA Flannery C.R., Little C.B.;
RT "Expression and activity of equine aggrecaenases.";
RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF368321; AK53425.1; -.
DR MEROPS: M12.221; -.
DR Interpro: IPR001590; Reprolysin.
DR PROSITE: PS00215; ADAM_MEPRO.1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 192
SQ SEQUENCE 192 AA: 20670 MW: 9013B0E19FCE8C56 CRC64:

```

```

Query Match 24.3%: Score 577; DB 6; Length 192;
Best Local Similarity 93.8%: Pred. No. 1.4e-36;
Matches 106; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

```

```

Qy 332 MADGVYCDPARSCAIVEDDGLQSAFTAHELGHVFNMLHNSKPCISLNGPLSTRHVA 391
Db 1 MADGVYCDPARSCAIVEDDGLQSAFTAHELGHVFNMLHNSKPCISLNGPLSTRHVA 60
Qy 392 APVMAHVDPPEPSPCSARITDFLDNGYGHCLDKPEAPLHPVYGVDDKDD 444
Db 61 APVMAHVDPPEPSPCSARITDFLDNGYGHCLDKPEAPLHPVYGVDDKDD 113

```

## RESULT 3

```

ID 019791 PRELIMINARY: PRT: 2165 AA.
AC 019791; Q27524;
DT 01-NOV-1996 (Tremblrel. 01, Created).
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE F25H8.3 protein.
GN F25H8.3.

```

```

OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Gajadaty S.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-94150718; PubMed-7906398;
RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latroille P.,
RA Lightning J., Lloyd C., Murray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy G., Ritken L., Roopra A., Saunders D., Shownkeen R.,
RA Smedon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Wellstock L., Wilkinson-Sproat J., Woldman P.;
RT *2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans *;
RL Nature 368:32-38(1994).
DR EMBL: 269361; CAA93288.1; -.
DR EMBL: 269360; CAA93288.1; JOINED.
DR EMBL: 269360; CAA93287.1; -.
DR EMBL: 269361; CAA93287.1; JOINED.
DR HSSP: P15167; IDTH.
DR MEROPS: M12.135; -.
DR Interpro: IPR001590; Reprolysin.
DR Interpro: IPR000884; TSP1.
DR Interpro: IPR00130; Zn_MTpeptidase.
DR Pfam: PF01421; Reprolysin; 1.
DR Pfam: PF00090; tsp.1; 14.
DR SMART: SM00209; TSP1; 18.
DR PROSITE: PS00215; ADAM_MEPRO.1.
DR PROSITE: PS00092; TSP1; 6.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 2165 AA: 244397 MW: FCC3DA8AAA9C4888 CRC64:

```

```

Query Match 22.4%: Score 532; DB 5; Length 2165;
Best Local Similarity 34.0%: Pred. No. 8.8e-32;
Matches 125; Conservative 61; Mismatches 126; Indels 56; Gaps 10;

```

```

Qy 109 VOYLGAPELLGS--AEPTGYLTGNG--DPESVSLHMDGALLGYOYRGAELHLP 164
Db 149 VLYLDEEEYRGMSRTDPCITRAHVKGVHSHIVLDCSEDLGYMLALPSGHTVER 208
Qy 165 LEGGTPNSAGPG--AHILRRKSP-----ASGCGPMCNVKAFL 200
Db 209 IISGNCTEHGASRHHOHLYRKFDPMHFKSFHLNSTSVNETETVATVADODGMEV---- 264
Qy 201 GSSPPRRPKRPA-SLSRVEVETLVYADDKMAAFHAGLKRILYLYMAAAAKAFKPSIR 259
Db 265 ---IERKASRRRAANSMDVVEVLYVADTKMYEYHGRSLIEDVYLTFSYVASIYRQHSR 321
Qy 260 NPVSLVTVRLVIGSGEGPOVPSAOTLRSCAMORGLNTPEDSDPHFDPAITFTRO 319
Db 322 ASINVVYVVKILYKTNAGPRTIIONAQTLODFCRMOQYTNDDPSVYHHDVAILLTKR 381
Qy 320 DLIC-GVSTCDTGLGADVGTCDPARSCAIVEDDGLQSAFTAHELGHVFNMLHNSKPC 378
Db 382 DICRSQGCDDTGLAELGTMCMDKSCAILEDNGLSAFTIAHELGHVSPIDHDKRC- 440
Qy 379 SLNGPLS-----TSRHVAPVMAHVDPPEPSPCSARITDFLDNGY- 421
Db 441 STYMPVNVKVKQSTKFDKTFQONNFHMAPLTLEYVTHPWSWSPCSAGMLERFLNNRGO 500
Qy 422 -HCLIDKP 428
Db 501 TQCLFDPP 508

```

RESULT 4  
 Q9W493 PRELIMINARY: PRT: 1054 AA.  
 ID Q9W493: 01-MAY-2000 (TREMBlrel. 13, Created)  
 AC 09W493: 01-MAY-2000 (TREMBlrel. 13, last sequence update)  
 DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)  
 DE CG4096 Protein.  
 GN CG4096.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxId=7227;  
 RX NCBI\_TaxId=7227;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blasej R.G., Champagne M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt A., Nelson C.R., Miklos G.L.G.,  
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktoglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,  
 RA Borova D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,  
 RA Burks K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durkin K.J., Evangelista C.C., Ferraz C., Fertler S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Galibert W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hosten D., Houston K.A., Howland T.J., Wei M.-H., Idegwam C.,  
 RA Jaisli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laso P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Maitel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merulov G., Mishina N.V., Modyarty C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,  
 RA Relbert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weisenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195 (2000).  
 DR EMBL: AE003435; AAF4605.1;  
 DR HSSP: P15167; IATL.  
 DR MEROPS: M12.231;  
 DR Flybase: FBgn0029791; CG4096.  
 DR InterPro: IPR002870; Pep\_M12B\_propep.  
 DR InterPro: IPR001590; ReprOlysin.  
 DR InterPro: IPR000884; TSPL.  
 DR InterPro: IPR000130; Zn\_MTpeptidse.  
 DR Pfam: PF01562; Pep\_M12B\_propep; 1.  
 DR Pfam: PF01421; ReprOlysin; 1.  
 DR Pfam: PF00090; tspl\_1; 2.  
 DR SMART: SM00209; TSPL; 2.  
 DR PROSITE: PS02025; ADAM\_MEPRO; 1.  
 DR PROSITE: PS00092; TSPL; 1.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; UNKNOWN; 1.  
 SO SEQUENCE 1054 AA; 118616 MW; DC15455555CB6212 CRC64;

Query Match 19.3%; Score 457; DB 5; Length 1054;  
 Best Local Similarity 43.7%; Pred. No. 1,9e-26;  
 Matches 100; Conservative 34; Mismatches 85; Indels 10; Gaps 7;  
 QY 207 PRARFASLSFVETLVVADKMAAFHAGLKRRLTVMAAARFAPHSIRNPVSLVY 266  
 DB 315 PRR-RRSISSPHVETLVVADKMAAFHAGLKRRLTVMAAARFAPHSIRNPVSLVY 372  
 QY 267 TRLVLGSGEEGPQVC--PSAQTLSRPFCAKRGINTPEPDSPPDHFTALITRQDLGV 324  
 DB 373 VRIQLDEEESQLQLNLQNAQKNLDRFCSWQKLNKSGEKDPHHDAVLLITRNIC-A 431  
 QY 325 STCDFTGMAVDGTCVDPARSCAIVEDDGLQSAFTAHELGHFNNMLDHSK-PCISLNGP 383  
 DB 432 NNMTLGLANVGKCKPKSCSVNEDNGIMLSHTITHELGHFGFHPDIAKICGHPRVGP 491  
 QY 384 LSTRRHVAPVMAHVDPEEPMSPCSARFTTDFLNGYGHCLLDKPEAPL 432  
 DB 492 IV---HIMPTFGADTLQVCMSCNSRKYTHFLDQGLGEC-LDDPPTPL 536

RESULT 5  
 Q8TE56 PRELIMINARY: PRT: 1095 AA.  
 ID Q8TE56: 01-JUN-2002 (TREMBlrel. 21, Created)  
 AC 08TE56: 01-JUN-2002 (TREMBlrel. 21, last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)  
 DE Metalloprotease disintegrin 17, with thrombospondin domains.  
 GN ADAMTS17.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euteria; Primates; Catarrhini; Homiidae; Homo.  
 NCBI\_TaxId=9606;  
 RX NCBI\_TaxId=9606;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21856482; PubMed=11867212;  
 RA Cal S., Obaaya A.J., Llamazares M., Garabaya C., Quesada V.,  
 RA Lopez-Otin C.;  
 RT "Cloning, expression analysis, and structural characterization of  
 RT seven novel human ADAMTS, a family of metalloproteinases with  
 RT disintegrin and thrombospondin-1 domains.";  
 RL Gene 283:49-62 (2002).  
 DR EMBL: AJ315735; CAC86016.1;  
 KW Integrin; Protease.  
 SO SEQUENCE 1095 AA; 121099 MW; A5C6540484754D5D CRC64;  
 Query Match 19.2%; Score 455; DB 4; Length 1095;  
 Best Local Similarity 32.2%; Pred. No. 2,9e-26;  
 Matches 149; Conservative 58; Mismatches 176; Indels 80; Gaps 18;  
 QY 22 PCLLFTVPLSLWMLLLLLASLPSARLASLPREEEL---VPEKLNGSVLPGSGT 77  
 DB 8 PPLVLPV-----LLLLWGLDPTGAVGAADAVEVLLPWRVPRDDVHLPPLPAAG 58  
 QY 78 PAR-----LLCRLOAFGETLLELBDSDSGVEGLTVQVIGQAPEL 118  
 DB 59 PRRRRPRTPPAPRAPRGERALLLHPAFGRDLVYQLKRDRLFSLRGFEVEEAGAAARR 118  
 QY 119 LGGAEPQTYLTGTINDPESVASLHWDG--GALLGVLYQGAELHLOPLEGCTPNSAGP 176  
 DB 119 GRAEALCFY--SGRVLHGPSGLVLSACGAAGLVGLIQGQEQVLIQPL----NNSGP 172  
 QY 177 GA---HILRRK-----SP-ASGGP--KCNV-----KAPLGSPSP--RRRAKRFSLSR 218  
 DB 173 FSGREHLIRKWSLTPSPSEAGRPOLCKVLTREKKRPGWRSDMRREARNAIRLTSR 232  
 QY 219 FVEETLVVADKMAAFHGA--GLKRYLLTVMAAARFAPHSIRNPVSLVTVRLVLSGEE 277  
 DB 233 TVETLVVADADWQYHGALEAQRFLITVNMNMVNMQHOSLGIKINIQVTKVLLRQRA 292  
 QY 278 GPVGVSAQOTLRSFCAMQ-----RGLNTPF-----DSDPDHFTTALITRQDL--GV 324

Db 293 KLSIGHHGESLSEFCHMOEEXEGARYLGNNOVPGKDDPPLVDAAVEFTRTDFCVHKD 352  
 325 STGDTGLMADVGTVCPARSCAIVEDDGLQSAFTAHAHELGHVFNMLHNDKPCISLNGPL 364  
 353 EPCDVTGIAVLGVCSSKRCVLAEDNGLATTTIAHELGHNGMKRHHDD-----HSSC 406  
 385 STSRHMAP--VMAHVDPPEWSPCSARFTTDPDNGYGHCLL 425  
 407 AGSHMSGEMVAKGRNPSDLSSSCSDLENFLEKSKVSTCLL 449

## RESULT 6

Q8TE60 PRELIMINARY: PRT: 1081 AA.

AC 08TE60: 01-JUN-2002 (TReMBLrel. 21, Created)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)  
 DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE ADAMTS18 protein.  
 GN ADAMTS18.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-21856482; PubMed-11867212;  
 RA Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V.,  
 RA Lopez-Otin C.,  
 RT "Cloning, expression analysis, and structural characterization of  
 RT seven novel human ADAMTSs, a family of metalloproteinases with  
 RT disintegrin and thrombospondin-1 domains.";  
 RT Gene 283:49-62(2002).  
 DR EMBL: AJ311903; CAC83612.1;  
 SO SEQUENCE 1081 AA; 119655 MW; 0438BF645676461E CRC64;

Query Match 17.7%; Score 418.5; DB 4; Length 1081;  
 Best local similarity 29.4%; Pred. No. 1.8e-23;  
 Matches 155; Conservative 55; Mismatches 196; Indels 121; Gaps 21;

QY 5 GSHRGGLA--GRLMGAPRCLLPIPLSLWLVLLLLLASLSPARLASPLPREBEY 62  
 Db 15 GSGPRGLGLGKRVAKALQCLC-----CASVAALASDSSSGASGLNDYVAV 64  
 QY 63 FPKRLN--GSVL-----PCSGTPARLLCQAEGFTLLLEEDSGVQEGTL 108  
 Db 65 TPEVDSAGSYISHDLHNGRKRKRSQNMARSSLIHYRFAFGDELHLEL-KPSILSSHFI 123  
 QY 109 VOYLG-----QAPELGCAEPGYLTGTINGDPESVASLHWGALLGVLOYGAELH 161  
 Db 124 VQVLGKDGASETKPREV-----OOCFYQGFIRNDSSSVAVSTCAG-LSGLIRTRNEFL 177  
 QY 162 LQPL-----EGCTPNSAGRGAILRRKS-----PASCQG----- 191  
 Db 178 ISPLPOLLAOEHNYSPPAGHNPVLYKTAEEKIQRYRGYPGSGRNYPGYSPSHIHASO 237  
 QY 192 -----RMCNVK--AP-----LGSPPRRRAKRFASLSREYE 221  
 Db 238 SRETEYHNRLQKHFCGRKKKAYARPTEDTYLRDEYGS-SGRRRRSAGSKQKLANE 296  
 QY 222 TLVVADDKMAAFGAG-LKRYLLTVMAAAAKAFKPSIRNPVSLVTVRIILGSGEGFO 280  
 Db 297 TLVVAADKKMEKKGKQNVTVTILTVKAVSG-LFKDGTIGSDINVVVVSLILLEQEBGILL 355  
 QY 281 VGSAAGOTLRSPCAMRGINTPEDSDPHFTALFTTRDLCGVST--CDTLGMADVGTV 338  
 Db 356 INHHADQSLNSFCQWASALIGKNGKRHDH--AILLTGDFISWKNEPCDTLGFAPTSGM 412  
 QY 339 CDPARSCAIVEDDGLQSAFTAHAHELGHVFNMLHD-NSKPCISLNGPLSTSRHMAPVMAH 397  
 Db 413 CSKVRSCITINEDTGLAFTIAHESGHNGMHDGEGNCRKRAEG-----NIMSFTLIG 466

QY 398 VDPEEPMSPCSARFTTDFLNGYGHCLLDPAPRLHPVTGYKDD 444  
 Db 467 NNGVFSWSSCSQYLKKFLSTPQAGCLYDPERK-----AGQVKYPD 507

## RESULT 7

Q8TE59 PRELIMINARY: PRT: 1207 AA.

AC 08TE59: 01-JUN-2002 (TReMBLrel. 21, Created)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)  
 DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE ADAMTS-19.  
 GN ADAMTS19.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-21856482; PubMed-11867212;  
 RA Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V.,  
 RA Lopez-Otin C.,  
 RT "Cloning, expression analysis, and structural characterization of  
 RT seven novel human ADAMTSs, a family of metalloproteinases with  
 RT disintegrin and thrombospondin-1 domains.";  
 RT Gene 283:49-62(2002).  
 DR EMBL: AJ311904; CAC84565.1;  
 SO SEQUENCE 1207 AA; 134061 MW; AF36F6BF5886FDE2 CRC64;

Query Match 16.6%; Score 394; DB 4; Length 1207;  
 Best local similarity 26.9%; Pred. No. 1.5e-21;  
 Matches 125; Conservative 62; Mismatches 194; Indels 84; Gaps 14;

QY 49 ARLASLPRE-----EEIVPEKNGSVLPSSGTPA----- 79  
 Db 101 SRLRPPPSGEEDELEQELPRGSSGAALSPGAPASMPPPPPPPPAQNAEP 160  
 QY 80 ---RLCRLQAFETLLLEEDSGVQEGTLVQY-----LGOAPELGCAEPGT 126  
 Db 161 DQDEVLLRIPAFSRDLYLLLRDGRFLARFAVEQRPNCPGTGAASAPQPPAPDACC 220  
 QY 127 YLTGTINGDPESVASLHWGALLGVLOYGAELHQLQPELGCTPNSAGRGAILRRK-- 184  
 Db 221 FYTGAVLRHPSGLASTGCGGIMGFITQINEDFIETPL-NDTMAITGHP-HRVYRQKSS 278  
 QY 185 -----SPASGCGPMQNVKAPLQSPSPR-----PRAKKRA-----SLSRFVETLVVADDKMA 231  
 Db 279 MEKVTESKALHSHYCGILSDKGRPPSRKIAESGRGKRYSYKLPQEVNIETVVVADPAMV 338  
 QY 232 AFHGA-GKRYLLTVMAAAAKAFKPSIRNPVSLVTVRIILGSGEGFOVPSAQTLR 290  
 Db 339 SYNGADAARFILLTILNMVFNLFQHKSLGVQVNLRYIKLILHETPRELYIGHGKMLE 398  
 QY 291 SFCAMOR-----GLNTPEDSDPHFTALFTTRDLC--GVSTCDTLGMADVT 337  
 Db 399 SFCMKQHEEFGKKNDIHLMSSTMGEDMTSVDAAILITRKDFVNHDECDTVGIAVLGSG 458  
 QY 338 VCDPARSCAIVEDDGLQSAFTAHAHELGHVFNMLHNDKPCISLNGPLSTSRHMAPVMAH 397  
 Db 459 KCSKKRCITIAEDNGLNLFTTIAHEGHNGMGINHDNDPSCACGLHIMSGEWIKGNLGD 518  
 QY 398 VDPEEPMSPCSARFTTDFLNGYGHCLLD-KPEA-----PLHLP 435  
 Db 519 VS-----WSKSKEDLERFLRSKASKNCLLOTNQSVASVAVPSKLP 559

## RESULT 8

Q8SXBO PRELIMINARY: PRT: 1688 AA.

AC 08SXBO: 01-JUN-2002 (TReMBLrel. 21, Created)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)



Best Local Similarity 86.2%; Pred. No. 6.3e-22;  
Matches 69; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

OY 365 HYFNHLHNSKRCISLUNGLSTRHVMAPVMAHVDPPEWMSCSARFTDFLNDNGYHCL 424  
|||||  
Db 1 HYFNHLHNSKRCISLUNGLSTRHVMAPVMAHVDPPEWMSCSARFTDFLNDNGYHCL 60  
|||||

OY 425 LDKPEAPLPLPTGDKDD 444  
|||||  
Db 61 LDKPEAPLPLPTGDKDD 80  
|||||

RESULT 11  
ID O9VF61 PRELIMINARY: PRT: 1229 AA.  
AC O9VF61:  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE CG6107 protein.  
GN CG6107.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Archopoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY.  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amaniatis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutcliffe G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazet R.G., Champagne M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhattacharya D., Bolshakov S.,  
RA Botkova D., Botchan M.A., Bouck J., Brokstein P., Brothier P.,  
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertirel W., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,  
RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mettel B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nussken D.R., Pacle J.M.,  
RA Palazzolo M., Plittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Sidenkimas I., Simpson M., Skupski M.P., Smith T.,  
RA Spletter E., Spreading A.C., Stapleton M., Strong R., Sun E.,  
RA Svrtkars R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster."  
RL Science 287:2185-2195(2000).  
DR EMBL: AE003709: AAF55199.1;  
DR FLYBase: FBgn0038340: CG6107.  
DR InterPro: IPR002870: Rep\_M12B\_propep.  
DR InterPro: IPR001590: RepPolysin.  
DR InterPro: IPR000884: TSP1.

DR InterPro: IPR00130: Zn\_Mrpeptidse.  
DR Pfam: PF01562: Rep\_M12B\_propep. 1.  
DR Pfam: PF01421: RepPolysin. 1.  
DR Pfam: PF00090: TSP1. 2.  
DR SMART: SM00209: TSP1. 3.  
DR PROSITE: PS02015: ADAM\_MPRO. 2.  
DR PROSITE: PS00092: TSP1. 2.  
DR PROSITE: PS00142: ZINC\_PROTEASE. 1.  
KW Hydrolase; Metalloprotease; Zinc.  
SQ SEQUENCE 1229 AA; 138136 MW; C1293160AB19B700 CRC64;

Query Match 14.9%; Score 353.5; DB 5; Length 1229;  
Best Local Similarity 28.3%; Pred. No. 2e-18;  
Matches 119; Conservative 60; Mismatches 131; Indels 111; Gaps 21;

OY 85 LQAFETLLEED-SCVOVEGLT-VOLG-----APELLGAEF---GTYLGTING 134  
|||  
Db 197 LNVFGRQLHLVLRDASFVHNSMTHIRLKGEEHPGEFEAEORHLLGCFYSGVED 256  
|||  
OY 135 DPESVASLHMDGALIGVLOYRGAELHLQPLEGTPNSAGPGAILRKRKSPASGCPMC 194  
|||  
Db 257 DPHSNVSVSLGG-WTGYIKTSFGALLQPVN---RTSSDEVLRHVRKSGRNARHVS 311  
|||  
OY 195 NKKAPLG-----SPSPRRRAKR-FASLSREY---ETLVVADDKMAAFHG 235  
|||  
Db 312 KEELGLDFMSKLEOVQEEBOKSKRLNKRKRHYADVQNVYTLLEVLIADVNSKQFHC 371  
|||  
OY 236 AGLKRYLLTYMAAAKAFKHPISLRPVSLVYRLVILSGEGCPQVP---SAAQTLSF 292  
|||  
Db 372 EDLQPIYLLIMSVISIPRDAISIGSIRLLVRLISL-----PINQDTHSSNMKHF 425  
|||  
OY 233 CAW--QRCLNPEDESDPHFDTAIFTRDGLGV---STCDTLGADVGYVCDPARSCAI 347  
|||  
Db 426 COFINSGER-----DTAMLTREPIGSGVPRKCHMLGLAELGTVCSSCSI 475  
|||  
OY 348 VEDDGLQSAFTAHLELGHVFNMLHNSKPCISLNGPLSTR-----HYMAPVMA-HWDP 400  
|||  
Db 476 VQDTLPLPAFTMAHLEIGHILNMHDDDKCM---PYVTRQNNKRVLIHIMSSVMGITHM-- 529  
|||  
OY 401 EEPMS-----PCARFTPLONGYHCLLD 426  
|||  
Db 530 -HPMSKTKDKSCLETSVGAHPIPTERLPEIISLDAQCOLSFGNDP-----GYCPFD 583  
|||  
OY 427 K 427  
|||  
Db 584 E 584  
|||

RESULT 12  
ID O8WXS8 PRELIMINARY: PRT: 1223 AA.  
AC O8WXS8:  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE A disintegrin-like and metalloprotease with thrombospondin type 1  
DE motif 14 precursor.  
GN ADAMTS14.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euteria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=21638061; PubMed=11779638;  
RA Boiz H., Ramirez A., von Bredehow B., Kubisch C.;  
RT "Characterization of ADAMTS14, a novel member of the ADAMTS  
RT metalloproteinase family."  
RL Biochim. Biophys. Acta 1522:221-225(2001).  
DR EMBL: AF358666: AAL40229.1;  
DR InterPro: IPR002870: Rep\_M12B\_propep.  
DR InterPro: IPR001590: RepPolysin.  
DR InterPro: IPR000884: TSP1.



QY 282 GPSAOTLSKFCAMQGLNTPEDSDPHFTAILFTRODLGVSCTDGLMADVTCVDP 341  
 Db 260 --NPSSTSLQVCMKMAHQROPSHAHHVFLTRDPF-GSSGNO--GYAPVTGKCHP 314  
 QY 342 ANSCAIVEDDGLQSAFTAAHELGHFVNMMLHNSKPCISLNGPLSTR--HYMAPVAHYD 399  
 Db 315 LNSCALNHDEGSSAFVIAHETGHVLMHEHDCG-----NGCADETSLSGVMAPLVQAFA 369  
 QY 400 PEEPSPSCARFTDFDLNCGYCHCLDKPEAP 431  
 Db 370 HRFHMSRCSKLESLRYLP--YDCLLDPPDP 399  
 RESULT 15  
 QY0126 PRELIMINARY: PRT: 1091 AA.  
 ID 09M126  
 AC 09M126; 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE CG3622 protein.  
 GN CG3622.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NC NCBI\_Taxid=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RC MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amandides P.G., Scherer S.E., Li P.W., Hoskins R.G., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.C., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Brill J.F., Abmayyari A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borovoy D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,  
 RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Rehert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stepleton M., Strong R., Sun E.,  
 RA Svitzkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weisslock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003458; AAF46905.1; -  
 DR Flybase: FBgn0034778; CG3622.  
 DR InterPro: IPR001590; Reprolysin.

DR InterPro: IPR001590; Zn\_MTPeptide.  
 DR Pfam: PF01421; Reprolysin. 1.  
 DR PROSITE: PSS0215; ADAM\_MPRO: 1.  
 DR PROSITE: PSS0142; ZINC\_PROTEASE; UNKNOWN.1.  
 SQ SEQUENCE 1091 AA: 124074 MW: 58967382F12565DE CRC64:  
 Query Match 13.2%; Score 312.5; DB 5; Length 1091;  
 Best local similarity 25.7%; Pred. No. 2.4e-15;  
 Matches 121; Conservative 59; Mismatches 179; Indels 11; Gaps 18;  
 QY 51 LNSPLPREBEIYFPEKLSVLPGSGT-----ARLLCRLOAFGETLLLEEDSGVOYE 105  
 Db 49 LAQORPEEQRLERRRRPGAEQHGNGPHQRRORSIPSEDCNGPLRPEEBEAGEDEVE 108  
 QY 106 GLTVQYLGAPELLGSA-----EPGTLYTGTINTDDPSVSLH-----MD 145  
 Db 109 ARL--QLQOSPOLIDDAFTFIRRTENGTOP--VERSPQLKRLERCFTSPAAALDCE 163  
 QY 146 GGALLGVLYRGAELHLOPLEGCTPNSAGPGAHIL-----RRKSPA 187  
 Db 164 PGNVRGVFOQNGSDLVIHPLP-----ARFGTGHVLYQARVDSGGGYSBASRRASP 218  
 QY 188 SGG-----GPKCNKAPLGSPP-----RR-----RAKRF--ASLSR 218  
 Db 219 DSOLOPEPDAEEFNEPERRRLRAOSQLOSPRLRFRORHHHPHQHQRRRYIGDAPRSK 278  
 QY 219 -----FVETLVVADDKMAF-----HGAGLKRLLTVMAAARAKFKPSIRNPV 262  
 Db 279 WSDIPPELFLEIETAFVDSDLXALMOKNFPTNTESKVVSEFLAMINGVOLLIHPTIGRRI 338  
 QY 263 SLVTVRLVILGSGEGSPQVPSAQT--LNSFCAMQGLNTPEDSDPHFTAILFTRODL 321  
 Db 339 NFVLKRLLEIWKSWDPGLVRSRDVENYVNSFCWKOEKLPFSDADPLHDHALVLTGLDI 398  
 QY 322 C-----GVSCDTLGMADVGVCDPARSCAIVEDDGLQSAFTAAHELGHFVNMMLHNSKPC 377  
 Db 399 VTYDKKANSQVGMATVGMCTSIYCTINEKHPESVFVAHNEGHNLMGRHDAKE-- 456  
 QY 378 ISLNGPLSTRHYMAVMAHYDPEEPSPSCARFTDFDLNCGYCHCLDK 427  
 Db 457 ISCD-----PTMHINSPKLG--SGKVMWSKCSRTYLDLFDLPQAECLFLDR 500

Search completed: April 21, 2003, 14:47:24  
 Job time: 58.3163 secs



GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 14:42:31 ; Search time 26.6581 Seconds

(without alignments)  
1611.968 Million cell updates/sec

Title: US-10-050-200-8  
Perfect score: 2370  
Sequence: 1 MSQTGSHPRGRLAGRLMGA.....PEAPLHLPVYNGYKDDDDKG 447

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: PIR.73:\*  
2: PIR1:\*  
3: PIR2:\*  
4: PIR3:\*  
5: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2318	97.8	837	2 T00355	hypothetical prote
2	979	41.3	951	2 T00017	gene ADAMS-1 prote
3	532	22.4	2165	2 T21371	hypothetical prote
4	360.5	15.2	1205	2 T18517	procollagen N-endo
5	249.5	10.5	903	2 S60257	meltrin alpha - mo
6	248.5	10.5	407	2 S66260	metalloproteinase
7	236.5	10.0	480	1 A30065	triglycerin precursor
8	225.5	9.5	478	2 A43296	atrolysin E (EC 3.
9	224.5	9.5	414	1 HYRSAC	atrolysin C (EC 3.
10	220.5	9.3	414	2 S41609	atrolysin C (EC 3.
11	216.5	9.1	411	1 HRSNEA	fibrolyase (EC 3.4.
12	212.5	9.0	478	2 J01301	hemorrhagic protei
13	212	8.9	481	2 JC4342	fibronolytic prote
14	210.5	8.9	609	2 S55270	carboxylastatin p
15	208.5	8.8	478	2 JC4880	fibronolytic metal
16	208	8.8	481	2 S43125	triglycerin precursor
17	208	8.8	617	2 S48160	metalloproteinase
18	204.5	8.6	519	2 S48169	metalloproteinase
19	203.5	8.6	514	2 S41608	atrolysin B (EC 3.
20	202	8.5	571	2 S24789	jatrachagin C (EC 3.
21	201	8.5	419	2 S41607	atrolysin A (EC 3.
22	197.5	8.3	1444	2 T18856	angiogenesis inhib
23	196.5	8.3	616	2 A55796	ecardin precursor -
24	196.5	8.3	860	2 T16892	hypothetical prote
25	193.5	8.0	610	2 JC7530	vascular apoptosis
26	189	8.0	826	2 A60385	monocycle surface a
27	187	7.9	814	2 G02390	disintegrin-like m
28	181.5	7.7	550	2 T47158	hypothetical prote
29	168.5	7.1	952	2 T18900	disintegrin and me

30	167	7.0	200	2 S15111	hemorrhagic factor
31	167	7.0	203	2 A59421	metalloproteinase
32	166	7.0	429	2 A42972	coagulation factor
33	159	6.7	957	2 T15976	hypothetical prote
34	157	6.6	202	1 HYRSR	ruberylin (EC 3.4
35	155.5	6.6	202	1 HYRVH2	hemorrhagic protei
36	155	6.5	416	2 A37877	hemorrhagic protei
37	154.5	6.5	670	2 I65967	disintegrin-like m
38	152.5	6.4	789	2 S28259	androgen-regulated
39	149.5	6.3	524	2 S38539	disintegrin-like m
40	146.5	6.2	660	2 S71949	metalloproteinase
41	146	6.1	419	2 A59414	metalloproteinase
42	144	6.1	203	2 S46443	adamalysin (EC 3.4
43	143	6.0	202	2 JC2550	Ac1 metalloprotein
44	143	6.0	905	2 S55059	fertilin alpha-I -
45	141.5	6.0	825	2 S55060	fertilin alpha-II

## ALIGNMENTS

## RESULT 1

T00355  
hypothetical protein KIAA0688 - human  
C:Species: Homo sapiens (man)  
C:Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 21-Jul-2000  
C:Accession: T00355  
R:Shikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, DNA Res. 5, 169-176, 1998  
A>Title: Prediction of the coding sequences of unidentified human genes. X. The compi  
A:Reference number: Z14142; MUID:98403880; PMID:9734811  
A:Accession: T00355  
A>Status: preliminary: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-837 <158>  
A:Cross-references: EMBL:AB014588; NID:93327189; PIDN:BAA31663.1; PID:93327190  
A:Experimental source: brain  
C:Genetics:  
A:Gene: KIAA0688  
C:Superfamily: thrombospondin type 1 repeat homology  
F:519-575/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match	97.8%;	Score 2318;	DB 2,	Length 837;
Best Local Similarity	99.1%;	Pred. No. 2e-172;		
Matches 440;	Conservative	0;	Mismatches 4;	Indels 0;
				Gaps 0;
QY	1	MSQTGSHPRGRLAGRLMGAQPCLLPIVPLSLVWLLLLLLASLPSARLASPLPREEE	60	
Db	1	MSQTGSHPRGRLAGRLMGAQPCLLPIVPLSLVWLLLLLLASLPSARLASPLPREEE	60	
QY	61	IVPEKLNQSVLPQSGTPARLLCRQAQFETLLLEEDSGVQVEGTIVYVGOAPELLG	120	
Db	61	IVPEKLNQSVLPQSGTPARLLCRQAQFETLLLEEDSGVQVEGTIVYVGOAPELLG	120	
QY	121	GAEPGYLTGTNGDPESVASHMDGALLGVLOYGAEHLQPLEGCTPNSAGCGAHI	180	
Db	121	GAEPGYLTGTNGDPESVASHMDGALLGVLOYGAEHLQPLEGCTPNSAGCGAHI	180	
QY	181	LRRKSPASGCGMPCNVKAPLVGSPSPRRRAKRFASLSRVEVLVYVADDMKMAAFHAGLKR	240	
Db	181	LRRKSPASGCGMPCNVKAPLVGSPSPRRRAKRFASLSRVEVLVYVADDMKMAAFHAGLKR	240	
QY	241	YLTVAANAARAKFKHPSINPVSILVYTRVILVIGSGEGPQVGSAAQOTLRSCAMORGLN	300	
Db	241	YLTVAANAARAKFKHPSINPVSILVYTRVILVIGSGEGPQVGSAAQOTLRSCAMORGLN	300	
QY	301	TPEDSPDHFDVAILFTRODLGCVSTCDTILGADVGTVCDDPARSCAIVDDDLQSAFTAA	360	
Db	301	TPEDSPDHFDVAILFTRODLGCVSTCDTILGADVGTVCDDPARSCAIVDDDLQSAFTAA	360	
QY	361	HELGHVFNMLHNSKPCISLNGPLSTSRVMAVYAHVPEEPWSPSCARFTTDFLDNGY	420	
Db	361	HELGHVFNMLHNSKPCISLNGPLSTSRVMAVYAHVPEEPWSPSCARFTTDFLDNGY	420	



C:Function:  
A:Description: catalyzes cleavage of the propeptides of type I and II collagens prior to  
C:Keywords: hydrolase; metalloproteinase

Query Match 15.28; Score 360.5; DB 2; Length 1205;  
Best Local Similarity 27.58; Pred. No. 9.1e-20;  
Matches 138; Conservative 65; Mismatches 178; Indels 121; Gaps 24;

QY 8 PGRGAGRWLGAQPCILLPIVPLSLWLLLLLLSLPSARLSPLP-----R 57  
DB 3 PPAAGAG-----LLCPA-----LLLLLLPLPADARLMAAADAAPPGPGQHGCA 46  
QY 58 EEETVFP-----EKLNGSVLPGSGTPA-----RLLCRLO 86  
DB 47 ERLIAPVPTDAGQRLVSHVSAATAPAGVTRBRAAPQITGLSGSEDEPGRLFYVVT 106  
QY 87 AFGELILLEDDSGVOVEGLTVQYLGAPPELLGCA---EP--GTYL--TGTING--DDES 138  
DB 107 VFGRLHLRLRPMARLVAPGATVEMOGES-----GATVEPLTCTLYGVDVAGLAESS 161  
QY 139 VASLHMDGALGLVQYRGAEHLQPLEGTPNSAGPG--AHILRRKSPASGCGPMCNVK 197  
DB 162 VALSMCDG--LAGILRMEEFETPELEKGLAAKEGGRVHVYHR--PTTSRPP----- 213  
QY 198 APLGSP-----SPRRARRFASLSRFVETLVVADKM 230  
DB 214 -PLGGPQALDTGISADSLDSLRALGVLEERVNSRRMRRAADDVNIIEVLGVDOSV 272  
QY 221 AAFHGA-GLKRYYLLTVMAAAKAFKHPISIRPVSLYTRLYTLGSGECP--QVGPSSAQ 287  
DB 273 VQFGEHVOYKYLTLTLMIVETIHDESLGAINVLAIRIILLSTGKSMLEIEIG-NSQ 331  
QY 288 TLRSFCAMQRLNTPEDSDPDHPTAIIETRODLCGVTCDTLGMADVGTVCDDPARSCAI 347  
DB 332 SLENNCRAYIQQKRPDTHDEYHDAIFLTRDP--GPGMG--GYAPYTGCHPFRSCTL 388  
QY 348 VEDDGLGSAFTAAHELGHVFNMLHD-NSKPCISLNGPLSTRHVPAPMAHVDEPEPSP 406  
DB 389 NHEDDFSSAFVAAHETGVHLMGHENDGQGNRC---GDEVRLGSIWAPVQAARFHFMSR 444  
QY 407 CSARITFDLNGYGHCLLDKP 428  
DB 445 CSQQLSRYL-HSY-DCLRDP 464

## RESULT 5

60257  
C:Species: Mus musculus (house mouse)  
C:Date: 10-Apr-1996 #sequence\_revision 19-Apr-1996 #text\_change 20-Jun-2000  
C:Accession: S60257  
R:Yagami-Hiromasa, T.; Sato, T.; Kurisaki, T.; Kamijo, K.; Nabeshima, Y.I.; Fujisawa-Set  
Nature 377, 655-656, 1995  
A:Title: A metalloproteinase-disintegrin participating in myoblast fusion.  
A:Reference number: S60257; MUID:96026308; PMID:7566181  
A:Accession: S60257  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-903 <YAG>  
A:Cross-references: EMBL:D50411; NID:91054586; PIDN:BA08912.1; PID:91054587  
C:Superfamily: mouse meltrin alpha; disintegrin homology  
F:421-503/domain: disintegrin homology <DIS>  
F:349/Active site: Glu #status predicted

Query Match 10.5%; Score 249.5; DB 2; Length 903;  
Best Local Similarity 25.7%; Pred. No. 2.8e-11;  
Matches 108; Conservative 62; Mismatches 203; Indels 47; Gaps 14;

QY 38 LLLALLSLPSARLASPPREEEIVFPEKINGSVL-----PGSGTPAR-----LLCRLO 86  
DB 15 LLLAALAGLLARARAGSLMDQRCATVEARASLLSKDPGIFGCSITPAKDHVDVTLQLO 74  
QY 87 AFGETILLEDDSGVOVEGLT-VQYL--GOAPPELLGGAEPCTYLTGTINGDPESVASLH 143

DB 75 LESRLILSLERNGLIANGFTEHYLQDGTDVSTTRNHTDHCYHHGVGDASVSL 134  
QY 144 WDGAALGLVQYRGAEHLQPLEGTPNSAGPGCAHILRRKSPASGCGPMCNVKAPLQSP 203  
DB 135 -TCSDLRLILFENKTYSLLEPMKNTDSYKLVPAESMNIQGLCSQHNKSLNLMEDVP 193  
QY 204 SPREPRAKRF-----ASLRFVETLVVADDMKMAAFGACL---KRYLLTVMAAAKAFKHP 256  
DB 194 GTSQMRARRHRREPLTKMYVELVYVADNREFOROKLEVKORLIEIANHVDFYR-- 251  
QY 257 SIRNPVSLVTRLYLG-----SGEGRQVGPSSAQTLLRSFCAMQRLNTPEDSDPHPT 312  
DB 252 ----PLNI---RIVLVEVEVNNIDDKCSISODPFLTHEFLDMWRIKLLPRKS-----HDN 300  
QY 313 AIIETRODLCGVSTCDTLGMADVGTVCDPARSCAIV---EDDGLQSAFTAAHELGHVNM 369  
DB 301 AOLISGVYFGST---TGMAPIMSMCTAEOGCGVMDHSDSPGAAVTLAHELGHNKG 356  
QY 370 LHDNSKPCISLNGPLSTRHVPAPMAHVDEPEPSPCSARITFDLNGYGHCLLDKPE 429  
DB 357 NHDILERCSCRMMAEKGGCINPSTGFPF--VVFSSCSRKDLASLEKMGKMLFNLPE 415

## RESULT 6

56260  
C:Species: Aphisrodon confortilx confortilx (southern copperhead)  
C:Date: 17-Dec-1996 #sequence\_revision 13-Mar-1997 #text\_change 22-Jun-1999  
C:Accession: S66260; S74263  
R:Selistre de Araujo, H.S.; Ombuy, C.L.  
Arch. Biochem. Biophys. 320, 141-148, 1995  
A:Title: Molecular cloning and sequence analysis of cDNAs for metalloproteinases from  
A:Reference number: S66260; MUID:95314311; PMID:7793974  
A:Accession: S66260  
A:Molecule type: mRNA  
A:Residues: 1-407 <SEL>  
A:Cross-references: EMBL:U18234; NID:9603216; PIDN:AAC59704.1; PID:9603217  
A:Accession: S74263  
A:Molecule type: protein  
A:Residues: 186-206 <SEA>  
C:Superfamily: atropisin C  
C:Keywords: hydrolase; metalloproteinase; zinc  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-187/Domain: propeptide #status predicted <PRO>  
F:188-407/Product: metalloproteinase #status experimental <MAT>  
F:304-384,344-351,346-366/Disulfide bonds: #status predicted  
F:329,333,339/Binding site: zinc, catalytic (His) #status predicted  
F:330/Active site: Glu #status predicted

Query Match 10.5%; Score 248.5; DB 2; Length 407;  
Best Local Similarity 26.4%; Pred. No. 1.2e-11;  
Matches 115; Conservative 72; Mismatches 178; Indels 71; Gaps 21;

QY 37 LLLALLSLP-----SARLASPLREEEIVPEKL-----NSVLPGSGTPARLLCRLOAF 88  
DB 5 LVLVTLCLAAFPYQGSIILESQNVNDYEVVYPRKYTPPKAGVQYKDYDAQY--ELKVN 62  
QY 89 GETLLELEDDSGVOVEGLTVQYL--GOAPPELLGGAEPCTYLTGTINGDPESVASLMD 145  
DB 63 GEPVVLHLERNKNGLPSKDYSETHYSPDRKRTTYPVPVDHCYHGRIONDASISASAC 122  
QY 146 GGALLGVQYRGAEHLQPLEGTPNSAGPGCAHILRRKSPASGCG---PMCNVAPLGS 202  
DB 123 NC-LGCHKRLQCEMLIEPLE-----LSDSRAHNVFTEYENVEKDEAPKICGVTQNMES 175  
QY 203 PSPRRARRKFPASLS-----RFVETLVVADDMK-AAFHGAG--LKRYYLLTVMAAAKAF 253  
DB 176 YEP-----IKKASQNLNLYOYQRYVELVTVVDHGMVTKXNGSDKLRQVWHQVMTMKESY 231  
QY 254 KHPISIRNPVSLVTRLYLGSGEGRQVGPSSAQTLLRSFCAMQRLNTPEDSDPH---F 310  
DB 232 RMYI--DISLNG---VEIWSNKDLIDVQPARHRLDSFGEMR-----ERDLHLRISH 279



Db 284 LRTSIAFDEQ-----IIGRAVIGICDPKRSSTGVODHSEINLRVAVTMTHELCHN 334  
 QY 367 FNMHLHD-NSKPCISLNGPLSTSRHMAVMAHVDEEPMSPCSARFTIDFNDNGYHCLL 425  
 Db 335 LGHNHDHSDSCSC-----GGYSC---IMSPVSD-EPSKIFSDCSYICQWETIMNOKPCITL 386  
 QY 426 DKP--EAPLHPVTGD 439  
 Db 387 KKPLRTDTVPSTVSGN 402

## RESULT 9

HYRSC  
 A:Allysin C (EC 3.4.24.42) precursor - western diamondback rattlesnake  
 N:Alternate names: Crotalus atrox metalloendopeptidase c; hemorrhagic toxin c; hemorrhagic  
 C:Species: Crotalus atrox (Western diamondback rattlesnake)  
 C>Date: 30-Sep-1992 #sequence\_revision 12-Apr-1996 #text\_change 18-Jun-1999  
 C:Accession: S41610; A34166  
 R:Hit: L.A.; Jla; L.G.; Bjarnason, J.B.; Fox, J.W.  
 Arch. Biochem. Biophys. 308, 182-191, 1994  
 A:Title: cDNA sequences for four snake venom metalloproteinases: structure, classifica  
 A:Reference number: S41607; MUID:94145078; PMID:8311451  
 A:Accession: S41610  
 A:Status: preliminary; translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-414 <HIT>  
 A:Cross-references: GB:U01227; NID:9402263; PIDN:AAA03352.1; PID:9402264  
 R:Shannon, J.D.; Barmova, E.N.; Bjarnason, J.B.; Fox, J.W.  
 J. Biol. Chem. 264, 11575-11583, 1989  
 A:Title: Amino acid sequence of a Crotalus atrox venom metalloproteinase which cleaves  
 A:Reference number: A34166; MUID:89308543; PMID:2745407  
 A:Accession: A34166  
 A:Molecule type: protein  
 A:Residues: 191-393 <SHA>  
 A:Note: 204-Ile was also found  
 A:Note: a variant lacking 192-Gln was also found  
 C:Superfamily: atrolysin C  
 C:Keywords: hydrolase; metalloproteinase; pyroglutamic acid; venom; zinc  
 F:1-12/Domain: signal sequence #status predicted <SIG>  
 F:13-19/Domain: amino-terminal propeptide #status predicted <PRO>  
 F:191-393/Product: atrolysin C #status experimental <MNT>  
 F:394-414/Domain: carboxyl-terminal propeptide #status predicted <CPT>  
 F:191/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experime  
 F:308-388,348-355/Disulfide bonds: #status experimental  
 F:333,337,343/Binding site: zinc (His) #status predicted  
 F:334/Active site: Glu #status predicted

Query Match 9.5% Score 224.5; DB 1; Length 414;  
 Best Local Similarity 26.4% Pred. No. 8.9e-10;  
 Matches 116; Conservative 64; Mismatches 186; Indels 73; Gaps 22;

QY 37 LLLLLLTLASLP-----SARLASPLPREEIVPEPKLNGSLP--GSGTPA---RLLCRLQAF 88  
 Db 5 LVVITICLAVFPYQSSITLESQNVNDYEVVPRKV--TALPKGAQVPRYEDAMQYELVYN 62  
 QY 89 GETLLLELEDSDSGVOVEGLTYOYL---GOAPELLGGAEPGYTLTGTINGDPESVASLHMD 145  
 Db 63 GEVYVLIHEKKKELESKDYSETHYSPDGRKITTPNSVEDHCYRRIENDADSTASISAC 122  
 QY 146 GGALLGVLOYRGAEHLHLPLEGCTPNSAGCGPANH--LRKSPASGGGPMCNVAPLGS 202  
 Db 123 NG-LKGHRKLOGEMLYIEPLE-----LSDSEAHAVFKLENVEKDEAPKMGVYQNMES 175  
 QY 203 PSPRRRAKRFASIS-----RFEVETLVVADDKMAAFHAGLKRKYLLTYMAAAKA 252  
 Db 176 YEP-----IKKASDLNLNDQONLPQRYELVYVADHRAVFMKNSDLN---TIRRVHEI 227  
 QY 253 FK-----HPSIRNPVSLVYTRLVILGSGEEGPQVPSAQTLSRCAMORG--LNTPEDS 305  
 Db 228 VNFINGFYRSLNIHVSLL--TDLEI--WSNEDQINIQSASDITLNAFAEREDLLNRKS-- 282  
 QY 306 DPHHEDTALITFRQDLGCVSTCDITLGMADVGTVCDPARSCAIVEDD---GQSATAAHE 362

Db 283 ----HDNAQLTATEL---DEETLGAPLCTMCDPKLSIGIVODHSEINLMLGVTMAHE 334  
 QY 363 LGHVFNMHLHDNSKPCISLNGPLSTSRHMAVMAHVDEEPMSPCSARFTIDFNDNGYH 422  
 Db 335 LGHNIGMEHD-GKDCLR-----GASLCIMRGLTKGNSTY-FSDDSMHYERFLKQYKPO 387  
 QY 423 CILDKPEA--PLHLHPVTGD 439  
 Db 388 CILNKLRLIDPVTSTVSGN 406

## RESULT 10

S41609  
 A:Allysin C (EC 3.4.24.42) precursor (Ht-c) - western diamondback rattlesnake  
 N:Alternate names: hemorrhagic toxin b  
 C:Species: Crotalus atrox (Western diamondback rattlesnake)  
 C>Date: 29-Sep-1994 #sequence\_revision 26-May-1995 #text\_change 22-Jun-1999  
 C:Accession: S41609  
 R:Hit: L.A.; Jla; L.G.; Bjarnason, J.B.; Fox, J.W.  
 Arch. Biochem. Biophys. 308, 182-191, 1994  
 A:Title: cDNA sequences for four snake venom metalloproteinases: structure, classific  
 A:Reference number: S41607; MUID:94145078; PMID:8311451  
 A:Accession: S41609  
 A:Status: preliminary; translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-414 <HIT>  
 A:Cross-references: EMBL:U01236; NID:9402261; PIDN:AAA03328.1; PID:9402262  
 C:Superfamily: atrolysin C  
 C:Keywords: hydrolase; metalloproteinase; pyroglutamic acid; venom; zinc  
 F:1-12/Domain: signal sequence #status predicted <SIG>  
 F:13-19/Domain: amino-terminal propeptide #status predicted <PRO>  
 F:191-393/Product: atrolysin C #status experimental <MNT>  
 F:394-414/Domain: carboxyl-terminal propeptide #status predicted <CPT>  
 F:191/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status exper  
 F:308-388,348-355/Disulfide bonds: #status predicted  
 F:333,337,343/Binding site: zinc (His) #status predicted  
 F:334/Active site: Glu #status predicted

Query Match 9.3% Score 220.5; DB 2; Length 414;  
 Best Local Similarity 26.2% Pred. No. 1.8e-09;  
 Matches 115; Conservative 65; Mismatches 186; Indels 73; Gaps 22;

QY 37 LLLLLLTLASLP-----SARLASPLPREEIVPEPKLNGSLP--GSGTPA---RLLCRLQAF 88  
 Db 5 LVVITICLAVFPYQSSITLESQNVNDYEVVPRKV--TALPKGAQVPRYEDAMQYELVYN 62  
 QY 89 GETLLLELEDSDSGVOVEGLTYOYL---GOAPELLGGAEPGYTLTGTINGDPESVASLHMD 145  
 Db 63 GEVYVLIHEKKKELESKDYSETHYSPDGRKITTPNSVEDHCYRRIENDADSTASISAC 122  
 QY 146 GGALLGVLOYRGAEHLHLPLEGCTPNSAGCGPANH--LRKSPASGGGPMCNVAPLGS 202  
 Db 123 NG-LKGHRKLOGEMLYIEPLE-----LSDSEAHAVFKLENVEKDEAPKMGVYQNMES 175  
 QY 203 PSPRRRAKRFASIS-----RFEVETLVVADDKMAAFHAGLKRKYLLTYMAAAKA 252  
 Db 176 YEP-----IKKASDLNLNDQONLPQRYELVYVADHRAVFMKNSDLN---TIRRVHEI 227  
 QY 253 FK-----HPSIRNPVSLVYTRLVILGSGEEGPQVPSAQTLSRCAMORG--LNTPEDS 305  
 Db 228 VNFINGFYRSLNIHVSLL--TDLEI--WSNEDQINIQSASDITLNAFAEREDLLNRKS-- 282  
 QY 306 DPHHEDTALITFRQDLGCVSTCDITLGMADVGTVCDPARSCAIVEDD---GQSATAAHE 362  
 Db 283 ----HDNAQLTATEL---DEETLGAPLCTMCDPKLSIGIVODHSEINLMLGVTMAHE 334  
 QY 363 LGHVFNMHLHDNSKPCISLNGPLSTSRHMAVMAHVDEEPMSPCSARFTIDFNDNGYH 422  
 Db 335 LGHNIGMEHD-GKDCLR-----GASLCIMRGLTKGNSTY-FSASMSHYERFLKQYKPO 387  
 QY 423 CILDKPEA--PLHLHPVTGD 439  
 Db 388 CILNKLRLIDPVTSTVSGN 406





Db 236 YRMYTH--VALGLE---TMSNEDKITVPEAGYTLNMGEMRK---TDLLTRKRH-DN 286  
OY 313 AILFTODLCGVSTCDTLGMAVGTVCDDPARSCAIVED---DGLQSAFTTAHFLGVFNM 369  
Db 287 AQLTLAIDLDLV-----IGLAVGSMCHRRKSTGIIDYSEINLVYAVIMAHMGHNLGI 341  
OY 370 LHDNSKPCISLNCPLSTSRH--VMAPVMAHVDEEPMSPCSARFTIDFDNGYGHCLDK 427  
Db 342 NHD-----SGYSCGQDYACIMRPEIS-PEPSTFFSNCSYECNDFIMNHNPECILNE 392  
OY 428 P 428  
Db 393 P 393

## RESULT 15

JC4880  
Fibrinolytic metalloproteinase (EC 3.4.24.-) precursor - Vipera lebetina  
N:Alternate names: lebetase Le3  
C:Species: Vipera lebetina  
C>Date: 10-Sep-1996 #sequence\_revision 18-Oct-1996 #text\_change 20-Jun-2000  
C:Accession: JC4880; PC4181  
R:Signif. E.: Aspollu, A.; Tu, A.T.; Sligur, J.  
Biochem. Biophys. Res. Commun. 224, 229-236, 1996  
A:Title: cDNA cloning and deduced amino acid sequence of fibrinolytic enzyme (lebetase)  
A:Reference number: JC4880; MUID:96280739; PMID:8694817  
A:Accession: JC4880  
A:Molecule type: mRNA  
A:Residues: 1-478 <SI11>  
A:Cross-references: EMBL:X97894; NID:g1502367; PIDN:CA66471.1; PID:g1502368  
A:Accession: PC4181  
A:Molecule type: protein  
A:Residues: 214-246 <SI12>  
A:Experimental source: venom  
C:Superfamily: trigramin precursor; disintegrin homology  
C:Keywords: anticoagulant; glycoprotein; hydrolase; metalloproteinase; venom; zinc  
F:1-18/Domain: signal sequence #status predicted <Sig>  
F:19-191/Domain: activation peptide #status predicted <ACP>  
F:192-397/Product: fibrinolytic metalloproteinase #status predicted <MAT>  
F:404-474/Domain: disintegrin homology <DIS>  
F:416-475/Domain: cell attachment (R-G-D) motif <DNT>  
F:456-458/Region: zinc attachment (Cys, His, His, His) (inhibited) #status  
F:170,337,341,347/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #status  
F:312-392,352-376,354-359/Disulfide bonds: #status predicted  
F:337,341,347/Binding site: zinc, catalytic (His) (active) #status predicted  
F:338/Active site: Glu #status predicted

Query Match 8 8%; Score 208.5; DB 2; Length 478;

Best Local Similarity 24.1%; Pred. No. 1.9e-08; Matches 107; Conservative 69; Mismatches 189; Indels 79; Gaps 20;

OY 37 LLLLLLASLP---SARLASPLPREEEIVFPEKLNGSVLPGSGTPAR-----LLCRLQA 87  
Db 5 LVTICLAVFPYGGSSKTLKSGNVNDYEVNPGAVTGLPKGANVQPEKKEEDTMQTEFEY 64  
OY 88 FGETLLELEQDSGVQVEGLT--VQYLGAPELL--CGAEPGYLVNGTINGDPESVASLHW 144  
Db 65 NGEPVVLHLEKNRGLFSKDYSETHYSPDGREITTNPAVEDHCYHGRIONDADSTASISA 124  
OY 145 DCGALLGVLDYRGAELHLPLEGSTPNSAGCPAHLIR--KSPASGQGP-MCNV----- 196  
Db 125 CNG-LKGYFTLRGETYLIEPLK--LPDSE---AAHVYKYENIEKDEAPKMGCVQTQNM 177  
OY 197 -----KAPLGSPSPRPRAKRFASLSRPVETLVVADDKMAAFHAGLKRYYLTVMAAA 249  
Db 178 ASDPEIKKASQNLTPEDQRF-----PRYIELYIVADHAMVTKYNGDL-----AA 223  
OY 250 AAKFKHPSIRN-----PVSIVTRLVYILSGEGEPQVGPSPAQTLRSPCAMQRLMT 301  
Db 224 ITTWVHQLVNNINGFYRLNVHTLSAVERWNTNGDL--INVQPAASVTLNLFCEWR----- 277  
OY 302 PEDSDPDHEDTALITFTRODLGVSTCDT--LGMAVGTVCDDPARSCAIVEDDGL---QSAF 357

Db 278 ---ERDLLNRRMHDAQLTLTGIDLDNIIIGLAYDDSCMDPRYSVGIYODHSAIIRLVAV 333  
OY 358 TAAHELGHVFNMLHDNSKPCISLNCPLSTSRHVMAPVMAHVDEEPMSPCSARFTIDFD 417  
Db 334 TAAHELGHNLGMNHDDQNCGANG-----CVMSVLLIE-QRSYQSPDCSKNKYQTYLT 386  
OY 418 NGYGHCLDKP--EAPLHLPTGSD 439  
Db 387 NRPQCILNQPRLRTDVTSTPVSQN 410

Search completed: April 21, 2003, 14:48:26  
Job time : 30.6581 secs





```
Db 121 GARPCTYLTCTINGDPEVSASLHMDGALLGVQYRGAEHLQPLEGCTPNSAGPGAH 180
Qy 181 LRRKSPASGCGPMCNKAPLGSPPRRRAKRFRASLSRFETLVVADDDKMAAFHAGLKR 240
Db 181 LRRKSPASGCGPMCNKAPLGSPPRRRAKRFRASLSRFETLVVADDDKMAAFHAGLKR 240
Qy 241 YLLTVMAAAAKAFKHPISIRNPVSLVTVRLVILSGEGEPGVGSAOQLRSFCAMQGLN 300
Db 241 YLLTVMAAAAKAFKHPISIRNPVSLVTVRLVILSGEGEPGVGSAOQLRSFCAMQGLN 300
Qy 301 TPEDSDPDHFDTAILEFTRODLGCVSTCDTLGMADVGVTCDPARSCAIVEDDGLQSAFTAA 360
Db 301 TPEDSDPDHFDTAILEFTRODLGCVSTCDTLGMADVGVTCDPARSCAIVEDDGLQSAFTAA 360
Qy 361 HELGHVFNMLHDSKPCISLNGPLSTSRHYMAPVMAHVDEEPMSPCSARFTIDFLDNGY 420
Db 361 HELGHVFNMLHDSKPCISLNGPLSTSRHYMAPVMAHVDEEPMSPCSARFTIDFLDNGY 420
Qy 421 GHCLLDKPEAPLHLPTVGYDKDD 444
Db 421 GHCLLDKPEAPLHLPTVGYDKDD 444
```

## RESULT 2

```
US-10-176-758-352
: Sequence 352, Application US/10176758
: Publication No. US2003008353A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Chen, Jian
: APPLICANT: Desnoyers, Luc
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Pan, James
: APPLICANT: Smith, Victoria
: APPLICANT: Watanabe, Colin K.
: APPLICANT: Wood, William I.
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P3430R1C104
: CURRENT APPLICATION NUMBER: US/10/176,758
: PRIOR FILING DATE: 2002-06-21
: PRIOR APPLICATION removed - See File Wrapper of Palm
: NUMBER OF SEQ ID NOS: 612
: SEQ ID NO 352
: LENGTH: 837
: TYPE: PRT
: ORGANISM: Homo Sapien
US-10-176-758-352
```

Query Match 97.3%; Score 2306; DB 9; Length 837;

Best Local Similarity 98.6%; Pred. No. 5, 6e-178; Matches 438; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy 1 MSOTGSHPGKGLAGRLMGAQPCLLPIVPLSWLWLLLLLLSLPSARLASPLPREEE 60
Db 1 MSOTGSHPGKGLAGRLMGAQPCLLPIVPLSWLWLLLLLLSLPSARLASPLPREEE 60
Qy 61 IYPERKLNLSVLPFGSGTPARLLCRLOAFGETLLLELEODSGVQVEGLTVQYLGQAPELLG 120
Db 61 IYPERKLNLSVLPFGSGTPARLLCRLOAFGETLLLELEODSGVQVEGLTVQYLGQAPELLG 120
Qy 121 GAEPGYTLTCTINGDPEVSASLHMDGALLGVQYRGAEHLQPLEGCTPNSAGPGAH 180
Db 121 GAEPGYTLTCTINGDPEVSASLHMDGALLGVQYRGAEHLQPLEGCTPNSAGPGAH 180
Qy 181 LRRKSPASGCGPMCNKAPLGSPPRRRAKRFRASLSRFETLVVADDDKMAAFHAGLKR 240
Db 181 LRRKSPASGCGPMCNKAPLGSPPRRRAKRFRASLSRFETLVVADDDKMAAFHAGLKR 240
```

```
Qy 241 YLLTVMAAAAKAFKHPISIRNPVSLVTVRLVILSGEGEPGVGSAOQLRSFCAMQGLN 300
Db 241 YLLTVMAAAAKAFKHPISIRNPVSLVTVRLVILSGEGEPGVGSAOQLRSFCAMQGLN 300
Qy 301 TPEDSDPDHFDTAILEFTRODLGCVSTCDTLGMADVGVTCDPARSCAIVEDDGLQSAFTAA 360
Db 301 TPEDSDPDHFDTAILEFTRODLGCVSTCDTLGMADVGVTCDPARSCAIVEDDGLQSAFTAA 360
Qy 361 HELGHVFNMLHDSKPCISLNGPLSTSRHYMAPVMAHVDEEPMSPCSARFTIDFLDNGY 420
Db 361 HELGHVFNMLHDSKPCISLNGPLSTSRHYMAPVMAHVDEEPMSPCSARFTIDFLDNGY 420
Qy 421 GHCLLDKPEAPLHLPTVGYDKDD 444
Db 421 GHCLLDKPEAPLHLPTVGYDKDD 444
```

## RESULT 3

```
US-10-175-737-352
: Sequence 352, Application US/10175737
: Publication No. US20030013153A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Chen, Jian
: APPLICANT: Desnoyers, Luc
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Pan, James
: APPLICANT: Smith, Victoria
: APPLICANT: Watanabe, Colin K.
: APPLICANT: Wood, William I.
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P3430R1C50
: CURRENT APPLICATION NUMBER: US/10/175,737
: PRIOR FILING DATE: 2002-06-19
: PRIOR APPLICATION removed - See File Wrapper of Palm
: NUMBER OF SEQ ID NOS: 612
: SEQ ID NO 352
: LENGTH: 837
: TYPE: PRT
: ORGANISM: Homo Sapien
US-10-175-737-352
```

Query Match 97.3%; Score 2306; DB 9; Length 837;

Best Local Similarity 98.6%; Pred. No. 5, 6e-178; Matches 438; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy 1 MSOTGSHPGKGLAGRLMGAQPCLLPIVPLSWLWLLLLLLSLPSARLASPLPREEE 60
Db 1 MSOTGSHPGKGLAGRLMGAQPCLLPIVPLSWLWLLLLLLSLPSARLASPLPREEE 60
Qy 61 IYPERKLNLSVLPFGSGTPARLLCRLOAFGETLLLELEODSGVQVEGLTVQYLGQAPELLG 120
Db 61 IYPERKLNLSVLPFGSGTPARLLCRLOAFGETLLLELEODSGVQVEGLTVQYLGQAPELLG 120
Qy 121 GAEPGYTLTCTINGDPEVSASLHMDGALLGVQYRGAEHLQPLEGCTPNSAGPGAH 180
Db 121 GAEPGYTLTCTINGDPEVSASLHMDGALLGVQYRGAEHLQPLEGCTPNSAGPGAH 180
Qy 181 LRRKSPASGCGPMCNKAPLGSPPRRRAKRFRASLSRFETLVVADDDKMAAFHAGLKR 240
Db 181 LRRKSPASGCGPMCNKAPLGSPPRRRAKRFRASLSRFETLVVADDDKMAAFHAGLKR 240
Qy 241 YLLTVMAAAAKAFKHPISIRNPVSLVTVRLVILSGEGEPGVGSAOQLRSFCAMQGLN 300
Db 241 YLLTVMAAAAKAFKHPISIRNPVSLVTVRLVILSGEGEPGVGSAOQLRSFCAMQGLN 300
Qy 301 TPEDSDPDHFDTAILEFTRODLGCVSTCDTLGMADVGVTCDPARSCAIVEDDGLQSAFTAA 360
Db 301 TPEDSDPDHFDTAILEFTRODLGCVSTCDTLGMADVGVTCDPARSCAIVEDDGLQSAFTAA 360
```

QY 361 HELGHVFNMLHDNSKPCISLNGPLSTSRHYMAPVMAHVDPEEPWSPCSARFTTDFLNGY 420  
DB 361 HELGHVFNMLHDNSKPCISLNGPLSTSRHYMAPVMAHVDPEEPWSPCSARFTTDFLNGY 420  
QY 421 GHCLLDKPEAPLHLPTVTDYKDD 444  
DB 421 GHCLLDKPEAPLHLPTVTDYKDD 444

## RESULT 4

US-10-173-706-352  
; Sequence 352, Application US/10173706  
; Publication No. US20030022293A1  
; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C7  
CURRENT APPLICATION NUMBER: US/10/173,706  
CURRENT FILING DATE: 2002-06-17  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 352  
LENGTH: 837  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-173-706-352

Query Match 97.3% Score 2306; DB 9; Length 837;  
Best Local Similarity 98.6% Pred. No. 5.6e-178;

Matches 438; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 MSQTSHPGRGLAGRWLMGAOPCLLPYVPLSVLWLLLLLLASLPSARLASPLPREEE 60  
DB 1 MSQTSHPGRGLAGRWLMGAOPCLLPYVPLSVLWLLLLLLASLPSARLASPLPREEE 60  
QY 61 IYPEKLNLSVLPDGSCTPARLLCRLQAFGETLLLEEDSGVOVEGLTVQYLGQAPELLG 120  
DB 61 IYPEKLNLSVLPDGSCTPARLLCRLQAFGETLLLEEDSGVOVEGLTVQYLGQAPELLG 120  
QY 121 GAEPGYLTGTINGDEPVSALHMDGALLGYQYRGAEHLQPLEGGTPNSAGGCAH 180  
DB 121 GAEPGYLTGTINGDEPVSALHMDGALLGYQYRGAEHLQPLEGGTPNSAGGCAH 180  
QY 181 LRKSPASGOGPMCNKAPLGSPSPRRARRKRFASLSREVELTVVADDDKMAAFHAGLKR 240  
DB 181 LRKSPASGOGPMCNKAPLGSPSPRRARRKRFASLSREVELTVVADDDKMAAFHAGLKR 240  
QY 241 YLTVAAMAAKAFKHPISIRNPVSLVYTRVLIIGSGEGEPQVGPSSAQTLSFCAMORGLN 300  
DB 241 YLTVAAMAAKAFKHPISIRNPVSLVYTRVLIIGSGEGEPQVGPSSAQTLSFCAMORGLN 300  
QY 301 TPEDSPDHPDFTAILFTRODLGVSCTDLGMADVTCVCDPARSCAIVEDDGQSFTAA 360  
DB 301 TPEDSPDHPDFTAILFTRODLGVSCTDLGMADVTCVCDPARSCAIVEDDGQSFTAA 360  
QY 361 HELGHVFNMLHDNSKPCISLNGPLSTSRHYMAPVMAHVDPEEPWSPCSARFTTDFLNGY 420  
DB 361 HELGHVFNMLHDNSKPCISLNGPLSTSRHYMAPVMAHVDPEEPWSPCSARFTTDFLNGY 420  
QY 421 GHCLLDKPEAPLHLPTVTDYKDD 444  
DB 421 GHCLLDKPEAPLHLPTVTDYKDD 444

DB 421 GHCLLDKPEAPLHLPTVTDYKDD 444

## RESULT 5

US-10-175-738-352  
; Sequence 352, Application US/10175738  
; Publication No. US20030022294A1  
; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C45  
CURRENT APPLICATION NUMBER: US/10/175,738  
CURRENT FILING DATE: 2002-06-19  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 352  
LENGTH: 837  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-175-738-352

Query Match 97.3% Score 2306; DB 9; Length 837;  
Best Local Similarity 98.6% Pred. No. 5.6e-178;  
Matches 438; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 MSQTSHPGRGLAGRWLMGAOPCLLPYVPLSVLWLLLLLLASLPSARLASPLPREEE 60  
DB 1 MSQTSHPGRGLAGRWLMGAOPCLLPYVPLSVLWLLLLLLASLPSARLASPLPREEE 60  
QY 61 IYPEKLNLSVLPDGSCTPARLLCRLQAFGETLLLEEDSGVOVEGLTVQYLGQAPELLG 120  
DB 61 IYPEKLNLSVLPDGSCTPARLLCRLQAFGETLLLEEDSGVOVEGLTVQYLGQAPELLG 120  
QY 121 GAEPGYLTGTINGDEPVSALHMDGALLGYQYRGAEHLQPLEGGTPNSAGGCAH 180  
DB 121 GAEPGYLTGTINGDEPVSALHMDGALLGYQYRGAEHLQPLEGGTPNSAGGCAH 180  
QY 181 LRKSPASGOGPMCNKAPLGSPSPRRARRKRFASLSREVELTVVADDDKMAAFHAGLKR 240  
DB 181 LRKSPASGOGPMCNKAPLGSPSPRRARRKRFASLSREVELTVVADDDKMAAFHAGLKR 240  
QY 241 YLTVAAMAAKAFKHPISIRNPVSLVYTRVLIIGSGEGEPQVGPSSAQTLSFCAMORGLN 300  
DB 241 YLTVAAMAAKAFKHPISIRNPVSLVYTRVLIIGSGEGEPQVGPSSAQTLSFCAMORGLN 300  
QY 301 TPEDSPDHPDFTAILFTRODLGVSCTDLGMADVTCVCDPARSCAIVEDDGQSFTAA 360  
DB 301 TPEDSPDHPDFTAILFTRODLGVSCTDLGMADVTCVCDPARSCAIVEDDGQSFTAA 360  
QY 361 HELGHVFNMLHDNSKPCISLNGPLSTSRHYMAPVMAHVDPEEPWSPCSARFTTDFLNGY 420  
DB 361 HELGHVFNMLHDNSKPCISLNGPLSTSRHYMAPVMAHVDPEEPWSPCSARFTTDFLNGY 420  
QY 421 GHCLLDKPEAPLHLPTVTDYKDD 444  
DB 421 GHCLLDKPEAPLHLPTVTDYKDD 444

## RESULT 6

US-10-175-752-352  
; Sequence 352, Application US/10175752  
; Publication No. US20030022295A1

```

: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Chen, Jian
: APPLICANT: Desnoyers, Luc
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Pan, James
: APPLICANT: Smith, Victoria
: APPLICANT: Watanabe, Colin K.
: APPLICANT: Wood, William I.
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P3430R1C60
: CURRENT APPLICATION NUMBER: US/10/175,752
: PRIOR APPLICATION: 2002-06-19
: NUMBER OF SEQ ID NOS: 612
: TYPE: PRF
: LENGTH: 837
: ORGANISM: Homo Sapien
US-10-175-752-352
```

Query Match 97.3% Score 2306; DB 9; Length 837;

Best Local Similarity 98.6% Pred. No. 5.6e-178; Matches 438; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

QY 1 MSOTGSHPGRGLAGRWLMGAPCLLPYPLSWLWLLLLASLPSARLASPLPREEB 60
DB 1 MSOTGSHPGRGLAGRWLMGAPCLLPYPLSWLWLLLLASLPSARLASPLPREEB 60
QY 61 IYPERKLSVLPVPGSGTPARLCRLQAFGETLLLEODSGVVEGLTYOYLQAPELLG 120
DB 61 IYPERKLSVLPVPGSGTPARLCRLQAFGETLLLEODSGVVEGLTYOYLQAPELLG 120
QY 121 GAEPGTYYLTGTINGDESVASLHMDGALLGVLYRGAEHLQPLEGCTPNSAGPGAH 180
DB 121 GAEPGTYYLTGTINGDESVASLHMDGALLGVLYRGAEHLQPLEGCTPNSAGPGAH 180
QY 181 LRRKSPASQSGPWCNKAPLGSPPRRARRKRFASLSRVEVTLYVADDDKMAAFHGGGLR 240
DB 181 LRRKSPASQSGPWCNKAPLGSPPRRARRKRFASLSRVEVTLYVADDDKMAAFHGGGLR 240
QY 241 YLTVAAMAAKAFKHPISINPVS LYVTVIILSGEGEGOVGSAOQLRSCFAMRGIN 300
DB 241 YLTVAAMAAKAFKHPISINPVS LYVTVIILSGEGEGOVGSAOQLRSCFAMRGIN 300
QY 301 TPEDSDPDHFDTAILETRDOLCGVSTCDTLGMADVCTVCDPARSCAIVEDDGLQSAFTA 360
DB 301 TPEDSDPDHFDTAILETRDOLCGVSTCDTLGMADVCTVCDPARSCAIVEDDGLQSAFTA 360
QY 361 HELGHVFNMLHDSKPCISLNGPLSTRHVMAVMAHVDPPEEWSPCSAFRTIDFLDNGY 420
DB 361 HELGHVFNMLHDSKPCISLNGPLSTRHVMAVMAHVDPPEEWSPCSAFRTIDFLDNGY 420
QY 421 GHCLLDKPEAPLPLPYTGDKDD 444
DB 421 GHCLLDKPEAPLPLPYTGDKDD 444
```

RESULT 7

US-10-176-482-352

Sequence 352, Application US/10176482

Publication No. US20030022296A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Chen, Jian

APPLICANT: Desnoyers, Luc

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Pan, James

APPLICANT: Smith, Victoria

APPLICANT: Watanabe, Colin K.

APPLICANT: Wood, William I.

```

: APPLICANT: Pan, James
: APPLICANT: Smith, Victoria
: APPLICANT: Watanabe, Colin K.
: APPLICANT: Wood, William I.
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P3430R1C70
: CURRENT APPLICATION NUMBER: US/10/176,482
: PRIOR APPLICATION: 2002-06-20
: NUMBER OF SEQ ID NOS: 612
: TYPE: PRF
: LENGTH: 837
: ORGANISM: Homo Sapien
US-10-176-482-352
```

Query Match 97.3% Score 2306; DB 9; Length 837;

Best Local Similarity 98.6% Pred. No. 5.6e-178; Matches 438; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

QY 1 MSOTGSHPGRGLAGRWLMGAPCLLPYPLSWLWLLLLASLPSARLASPLPREEB 60
DB 1 MSOTGSHPGRGLAGRWLMGAPCLLPYPLSWLWLLLLASLPSARLASPLPREEB 60
QY 61 IYPERKLSVLPVPGSGTPARLCRLQAFGETLLLEODSGVVEGLTYOYLQAPELLG 120
DB 61 IYPERKLSVLPVPGSGTPARLCRLQAFGETLLLEODSGVVEGLTYOYLQAPELLG 120
QY 121 GAEPGTYYLTGTINGDESVASLHMDGALLGVLYRGAEHLQPLEGCTPNSAGPGAH 180
DB 121 GAEPGTYYLTGTINGDESVASLHMDGALLGVLYRGAEHLQPLEGCTPNSAGPGAH 180
QY 181 LRRKSPASQSGPWCNKAPLGSPPRRARRKRFASLSRVEVTLYVADDDKMAAFHGGGLR 240
DB 181 LRRKSPASQSGPWCNKAPLGSPPRRARRKRFASLSRVEVTLYVADDDKMAAFHGGGLR 240
QY 241 YLTVAAMAAKAFKHPISINPVS LYVTVIILSGEGEGOVGSAOQLRSCFAMRGIN 300
DB 241 YLTVAAMAAKAFKHPISINPVS LYVTVIILSGEGEGOVGSAOQLRSCFAMRGIN 300
QY 301 TPEDSDPDHFDTAILETRDOLCGVSTCDTLGMADVCTVCDPARSCAIVEDDGLQSAFTA 360
DB 301 TPEDSDPDHFDTAILETRDOLCGVSTCDTLGMADVCTVCDPARSCAIVEDDGLQSAFTA 360
QY 361 HELGHVFNMLHDSKPCISLNGPLSTRHVMAVMAHVDPPEEWSPCSAFRTIDFLDNGY 420
DB 361 HELGHVFNMLHDSKPCISLNGPLSTRHVMAVMAHVDPPEEWSPCSAFRTIDFLDNGY 420
QY 421 GHCLLDKPEAPLPLPYTGDKDD 444
DB 421 GHCLLDKPEAPLPLPYTGDKDD 444
```

RESULT 8

US-10-176-757-352

Sequence 352, Application US/10176757

Publication No. US20030022297A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Chen, Jian

APPLICANT: Desnoyers, Luc

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Pan, James

APPLICANT: Smith, Victoria

APPLICANT: Watanabe, Colin K.

APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

```
FILE REFERENCE: P3430R1C86
CURRENT APPLICATION NUMBER: US/10/176,757
CURRENT FILING DATE: 2002-06-20
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 352
LENGTH: 837
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-757-352

Query Match
Best Local Similarity 97.3%; Score 2306; DB 9; Length 837;
Pred. No. 5,6e-178;
Matches 438; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 MSOTGSHPGKGLAGRMWGAOPCLLPYIPLSLWMLLLLLLLASLPSARLASPLPREEE 60
DB 1 MSOTGSHPGKGLAGRMWGAOPCLLPYIPLSLWMLLLLLLLASLPSARLASPLPREEE 60
QY 61 IYPERKLSVLPFGSGAPARLLCRLQAFGETLLLEEDSGVQVEGLTVQYLGAPELLG 120
DB 61 IYPERKLSVLPFGSGAPARLLCRLQAFGETLLLEEDSGVQVEGLTVQYLGAPELLG 120
QY 121 GAEPGYTLTGTTNGDEPESVSLHWDGALLGVQYRGAEHLLOPLRGCTPNSAGPGAH 180
DB 121 GAEPGYTLTGTTNGDEPESVSLHWDGALLGVQYRGAEHLLOPLRGCTPNSAGPGAH 180
QY 121 GAEPGYTLTGTTNGDEPESVSLHWDGALLGVQYRGAEHLLOPLRGCTPNSAGPGAH 180
DB 121 GAEPGYTLTGTTNGDEPESVSLHWDGALLGVQYRGAEHLLOPLRGCTPNSAGPGAH 180
QY 181 LRKRSASGCGPMCNKAPLGSPPRRRAKRFASLSRFVETLVVADDDKMAAFHAGLKR 240
DB 181 LRKRSASGCGPMCNKAPLGSPPRRRAKRFASLSRFVETLVVADDDKMAAFHAGLKR 240
QY 241 YLLTVMAAAAKAFKHPISIRNPVSLVYTRVLIIGSGEGPQVGPSSAQTLSFCAMORGLN 300
DB 241 YLLTVMAAAAKAFKHPISIRNPVSLVYTRVLIIGSGEGPQVGPSSAQTLSFCAMORGLN 300
QY 301 TPEDSDPDHFDTAIILFTRQDLGCVSTCDTLGMADVGTVCDPARSCAIVEDDGLQSAFTAA 360
DB 301 TPEDSDPDHFDTAIILFTRQDLGCVSTCDTLGMADVGTVCDPARSCAIVEDDGLQSAFTAA 360
QY 361 HELGHVFNMHDSKPCISLNGPLSTRHYMAPVMAHVDEEPMSPCSARFTIDFLDNGY 420
DB 361 HELGHVFNMHDSKPCISLNGPLSTRHYMAPVMAHVDEEPMSPCSARFTIDFLDNGY 420
QY 421 GHCLLDKPEAPLHLVPTGYDKDD 444
DB 421 GHCLLDKPEAPLHLVPTGYDKDD 444

RESULT 9
US-10-176-913-352
Sequence 352, Application US/10176913
Publication No. US20030022298A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C86
CURRENT APPLICATION NUMBER: US/10/176,913
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 352
LENGTH: 837
```

```
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-913-352

Query Match
Best Local Similarity 97.3%; Score 2306; DB 9; Length 837;
Pred. No. 5,6e-178;
Matches 438; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 MSOTGSHPGKGLAGRMWGAOPCLLPYIPLSLWMLLLLLLLASLPSARLASPLPREEE 60
DB 1 MSOTGSHPGKGLAGRMWGAOPCLLPYIPLSLWMLLLLLLLASLPSARLASPLPREEE 60
QY 61 IYPERKLSVLPFGSGAPARLLCRLQAFGETLLLEEDSGVQVEGLTVQYLGAPELLG 120
DB 61 IYPERKLSVLPFGSGAPARLLCRLQAFGETLLLEEDSGVQVEGLTVQYLGAPELLG 120
QY 121 GAEPGYTLTGTTNGDEPESVSLHWDGALLGVQYRGAEHLLOPLRGCTPNSAGPGAH 180
DB 121 GAEPGYTLTGTTNGDEPESVSLHWDGALLGVQYRGAEHLLOPLRGCTPNSAGPGAH 180
QY 121 GAEPGYTLTGTTNGDEPESVSLHWDGALLGVQYRGAEHLLOPLRGCTPNSAGPGAH 180
DB 121 GAEPGYTLTGTTNGDEPESVSLHWDGALLGVQYRGAEHLLOPLRGCTPNSAGPGAH 180
QY 181 LRKRSASGCGPMCNKAPLGSPPRRRAKRFASLSRFVETLVVADDDKMAAFHAGLKR 240
DB 181 LRKRSASGCGPMCNKAPLGSPPRRRAKRFASLSRFVETLVVADDDKMAAFHAGLKR 240
QY 241 YLLTVMAAAAKAFKHPISIRNPVSLVYTRVLIIGSGEGPQVGPSSAQTLSFCAMORGLN 300
DB 241 YLLTVMAAAAKAFKHPISIRNPVSLVYTRVLIIGSGEGPQVGPSSAQTLSFCAMORGLN 300
QY 301 TPEDSDPDHFDTAIILFTRQDLGCVSTCDTLGMADVGTVCDPARSCAIVEDDGLQSAFTAA 360
DB 301 TPEDSDPDHFDTAIILFTRQDLGCVSTCDTLGMADVGTVCDPARSCAIVEDDGLQSAFTAA 360
QY 361 HELGHVFNMHDSKPCISLNGPLSTRHYMAPVMAHVDEEPMSPCSARFTIDFLDNGY 420
DB 361 HELGHVFNMHDSKPCISLNGPLSTRHYMAPVMAHVDEEPMSPCSARFTIDFLDNGY 420
QY 421 GHCLLDKPEAPLHLVPTGYDKDD 444
DB 421 GHCLLDKPEAPLHLVPTGYDKDD 444

RESULT 10
US-10-180-552-352
Sequence 352, Application US/10180552
Publication No. US20030022300A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C153
CURRENT APPLICATION NUMBER: US/10/180,552
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 352
LENGTH: 837
TYPE: PRT
ORGANISM: Homo Sapien
US-10-180-552-352

Query Match
Best Local Similarity 97.3%; Score 2306; DB 9; Length 837;
Pred. No. 5,6e-178;
Matches 438; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```

QY 1 MSOTGSHPGRLAGRLWLMGAOPCLLLPIVPLSWLWLLLLLLASLPSARLASPLPREEE 60
DB 1 MSOTGSHPGRLAGRLWLMGAOPCLLLPIVPLSWLWLLLLLLASLPSARLASPLPREEE 60
QY 61 IYPERKINGSVLPGSGTPARLLCRLQAFGETLLLELEODSGVQVEGLTVQYLGQAPELLG 120
DB 61 IYPERKINGSVLPGSGTPARLLCRLQAFGETLLLELEODSGVQVEGLTVQYLGQAPELLG 120
QY 121 GAEPTGYLTGTINGDEPESVSLHMDGALLGVQYRGAEHLQPLEGGPNSAGGPAH 180
DB 121 GAEPTGYLTGTINGDEPESVSLHMDGALLGVQYRGAEHLQPLEGGPNSAGGPAH 180
QY 181 LRRKSPASGCGPMCNKAPLGSPSPRRRAKRFASLSRFVETLVVADDDKMAAFHAGLKR 240
DB 181 LRRKSPASGCGPMCNKAPLGSPSPRRRAKRFASLSRFVETLVVADDDKMAAFHAGLKR 240
QY 241 YLLTVMAAAAKAFKHPISIRNPVSLVYTRVLYLGSGECPGVGSAOQLRSCAMORGLN 300
DB 241 YLLTVMAAAAKAFKHPISIRNPVSLVYTRVLYLGSGECPGVGSAOQLRSCAMORGLN 300
QY 301 TPEDSDPDHFDTAILLFTRODLGCVSTCDTLGMADVGTCDPARSCAIVEDDGLQSAFTAA 360
DB 301 TPEDSDPDHFDTAILLFTRODLGCVSTCDTLGMADVGTCDPARSCAIVEDDGLQSAFTAA 360
QY 361 HELGHVFNMLHDNSKPCISLNGPLSTSRRHYMAFVMAHVDEEPMSPCSARFITDFLDNGY 420
DB 361 HELGHVFNMLHDNSKPCISLNGPLSTSRRHYMAFVMAHVDEEPMSPCSARFITDFLDNGY 420
QY 421 GHCLLDKPEAPLHLPVTGDKDD 444
DB 421 GHCLLDKPEAPLHLPVTGDKDD 444

```

## RESULT 11

```

US-10-180-557-352
: Sequence 352, Application US/10180557
: Publication No. US20030022301A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Chen, Jian
: APPLICANT: Desnoyers, Luc
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Pan, James
: APPLICANT: Smith, Victoria
: APPLICANT: Watanabe, Colin K.
: APPLICANT: Wood, William I.
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P3430R1C147
: CURRENT APPLICATION NUMBER: US/10/180,557
: Prior Application removed - 2002-06-25
: NUMBER OF SEQ ID NOS: 612
: SEQ ID NO 352
: LENGTH: 837
: TYPE: PRT
: ORGANISM: Homo Sapien
US-10-180-557-352

```

Query Match 97.3%; Score 2306; DB 9; Length 837;

Best Local Similarity 98.6%; Pred. No. 5,6e-178;

Matches 438; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

QY 1 MSOTGSHPGRLAGRLWLMGAOPCLLLPIVPLSWLWLLLLLLASLPSARLASPLPREEE 60
DB 1 MSOTGSHPGRLAGRLWLMGAOPCLLLPIVPLSWLWLLLLLLASLPSARLASPLPREEE 60
QY 61 IYPERKINGSVLPGSGTPARLLCRLQAFGETLLLELEODSGVQVEGLTVQYLGQAPELLG 120
DB 61 IYPERKINGSVLPGSGTPARLLCRLQAFGETLLLELEODSGVQVEGLTVQYLGQAPELLG 120

```

```

DB 61 IYPERKINGSVLPGSGTPARLLCRLQAFGETLLLELEODSGVQVEGLTVQYLGQAPELLG 120
QY 121 GAEPTGYLTGTINGDEPESVSLHMDGALLGVQYRGAEHLQPLEGGPNSAGGPAH 180
DB 121 GAEPTGYLTGTINGDEPESVSLHMDGALLGVQYRGAEHLQPLEGGPNSAGGPAH 180
QY 181 LRRKSPASGCGPMCNKAPLGSPSPRRRAKRFASLSRFVETLVVADDDKMAAFHAGLKR 240
DB 181 LRRKSPASGCGPMCNKAPLGSPSPRRRAKRFASLSRFVETLVVADDDKMAAFHAGLKR 240
QY 241 YLLTVMAAAAKAFKHPISIRNPVSLVYTRVLYLGSGECPGVGSAOQLRSCAMORGLN 300
DB 241 YLLTVMAAAAKAFKHPISIRNPVSLVYTRVLYLGSGECPGVGSAOQLRSCAMORGLN 300
QY 301 TPEDSDPDHFDTAILLFTRODLGCVSTCDTLGMADVGTCDPARSCAIVEDDGLQSAFTAA 360
DB 301 TPEDSDPDHFDTAILLFTRODLGCVSTCDTLGMADVGTCDPARSCAIVEDDGLQSAFTAA 360
QY 361 HELGHVFNMLHDNSKPCISLNGPLSTSRRHYMAFVMAHVDEEPMSPCSARFITDFLDNGY 420
DB 361 HELGHVFNMLHDNSKPCISLNGPLSTSRRHYMAFVMAHVDEEPMSPCSARFITDFLDNGY 420
QY 421 GHCLLDKPEAPLHLPVTGDKDD 444
DB 421 GHCLLDKPEAPLHLPVTGDKDD 444

```

## RESULT 12

```

US-10-173-700-352
: Sequence 352, Application US/10173700
: Publication No. US20030027262A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Chen, Jian
: APPLICANT: Desnoyers, Luc
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Pan, James
: APPLICANT: Smith, Victoria
: APPLICANT: Watanabe, Colin K.
: APPLICANT: Wood, William I.
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P3430R1C14
: CURRENT APPLICATION NUMBER: US/10/173,700
: Prior Application removed - See file wrapper or "Palm
: NUMBER OF SEQ ID NOS: 612
: SEQ ID NO 352
: LENGTH: 837
: TYPE: PRT
: ORGANISM: Homo Sapien
US-10-173-700-352

```

Query Match 97.3%; Score 2306; DB 9; Length 837;

Best Local Similarity 98.6%; Pred. No. 5,6e-178;

Matches 438; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

QY 1 MSOTGSHPGRLAGRLWLMGAOPCLLLPIVPLSWLWLLLLLLASLPSARLASPLPREEE 60
DB 1 MSOTGSHPGRLAGRLWLMGAOPCLLLPIVPLSWLWLLLLLLASLPSARLASPLPREEE 60
QY 61 IYPERKINGSVLPGSGTPARLLCRLQAFGETLLLELEODSGVQVEGLTVQYLGQAPELLG 120
DB 61 IYPERKINGSVLPGSGTPARLLCRLQAFGETLLLELEODSGVQVEGLTVQYLGQAPELLG 120
QY 121 GAEPTGYLTGTINGDEPESVSLHMDGALLGVQYRGAEHLQPLEGGPNSAGGPAH 180
DB 121 GAEPTGYLTGTINGDEPESVSLHMDGALLGVQYRGAEHLQPLEGGPNSAGGPAH 180
QY 181 LRRKSPASGCGPMCNKAPLGSPSPRRRAKRFASLSRFVETLVVADDDKMAAFHAGLKR 240
DB 181 LRRKSPASGCGPMCNKAPLGSPSPRRRAKRFASLSRFVETLVVADDDKMAAFHAGLKR 240

```

```

Db 181 LRRKSPASQCGPWCNKVAPLGSPPRRARRFASLSRFVETLVVADDKMAAFHGGGLKR 240
QY 241 YLLTVAATAAKAFKHPISINPVSIVYTRVYILSGGEGQVGPSPAQOTLRSCAMORGIN 300
Db 241 YLLTVAATAAKAFKHPISINPVSIVYTRVYILSGGEGQVGPSPAQOTLRSCAMORGIN 300
QY 301 TPEDSDPDHFDTAILEFTRODLCGVSTCDTLGNMADVGTVCDDPARSCAIVEDDGLQSAFTAA 360
Db 301 TPEDSGPDHFDTAILEFTRODLCGVSTCDTLGNMADVGTVCDDPARSCAIVEDDGLQSAFTAA 360
QY 361 HELGHVFNMLHDNSKPCISLNGPLSTRHVMAPVMAHVDPEEPWSPCSARFTTDFLDNGY 420
Db 361 HELGHVFNMLHDNSKPCISLNGPLSTRHVMAPVMAHVDPEEPWSPCSARFTTDFLDNGY 420
QY 421 GHCLLDKPEAPLHLPTGYDKDD 444
Db 421 GHCLLDKPEAPLHLPTGYDKDD 444

```

RESULT 13

```

US-10-174-572-352
; Sequence 352, Application US/10174572
; Publication No. US20030027263A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC40
; CURRENT APPLICATION NUMBER: US/10/174,572
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 352
; LENGTH: 837
; TYPE: PRP
; ORGANISM: Homo Sapien
US-10-174-572-352

```

Query Match 97.3%; Score 2306; DB 9; Length 837;  
 Best Local Similarity 98.6%; Pred. No. 5,6e-178;  
 Matches 438; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

QY 1 MSOTGSHPRGRLAGRWLMGAOPCLLPVPLSMVLWLLLLASLPSARLASPLPREE 60
Db 1 MSOTGSHPRGRLAGRWLMGAOPCLLPVPLSMVLWLLLLASLPSARLASPLPREE 60
QY 61 IYFPEKLSVLPDGSAPRLCLRLQAFGETLLLEEDSGVQVEGLTVQYLGQAPELLG 120
Db 61 IYFPEKLSVLPDGSAPRLCLRLQAFGETLLLEEDSGVQVEGLTVQYLGQAPELLG 120
QY 121 GAEPGYLTGTINGDESVASLHMDGALLGVLYRGAELHQLPLEGTPNSAGGGAHI 180
Db 121 GAEPGYLTGTINGDESVASLHMDGALLGVLYRGAELHQLPLEGTPNSAGGGAHI 180
QY 181 LRRKSPASQCGPWCNKVAPLGSPPRRARRFASLSRFVETLVVADDKMAAFHGGGLKR 240
Db 181 LRRKSPASQCGPWCNKVAPLGSPPRRARRFASLSRFVETLVVADDKMAAFHGGGLKR 240
QY 241 YLLTVAATAAKAFKHPISINPVSIVYTRVYILSGGEGQVGPSPAQOTLRSCAMORGIN 300
Db 241 YLLTVAATAAKAFKHPISINPVSIVYTRVYILSGGEGQVGPSPAQOTLRSCAMORGIN 300

```

```

QY 301 TPEDSDPDHFDTAILEFTRODLCGVSTCDTLGNMADVGTVCDDPARSCAIVEDDGLQSAFTAA 360
Db 301 TPEDSGPDHFDTAILEFTRODLCGVSTCDTLGNMADVGTVCDDPARSCAIVEDDGLQSAFTAA 360
QY 361 HELGHVFNMLHDNSKPCISLNGPLSTRHVMAPVMAHVDPEEPWSPCSARFTTDFLDNGY 420
Db 361 HELGHVFNMLHDNSKPCISLNGPLSTRHVMAPVMAHVDPEEPWSPCSARFTTDFLDNGY 420
QY 421 GHCLLDKPEAPLHLPTGYDKDD 444
Db 421 GHCLLDKPEAPLHLPTGYDKDD 444

```

RESULT 14

```

US-10-174-579-352
; Sequence 352, Application US/10174579
; Publication No. US20030027264A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC31
; CURRENT APPLICATION NUMBER: US/10/174,579
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 352
; LENGTH: 837
; TYPE: PRP
; ORGANISM: Homo Sapien
US-10-174-579-352

```

Query Match 97.3%; Score 2306; DB 9; Length 837;  
 Best Local Similarity 98.6%; Pred. No. 5,6e-178;  
 Matches 438; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

QY 1 MSOTGSHPRGRLAGRWLMGAOPCLLPVPLSMVLWLLLLASLPSARLASPLPREE 60
Db 1 MSOTGSHPRGRLAGRWLMGAOPCLLPVPLSMVLWLLLLASLPSARLASPLPREE 60
QY 61 IYFPEKLSVLPDGSAPRLCLRLQAFGETLLLEEDSGVQVEGLTVQYLGQAPELLG 120
Db 61 IYFPEKLSVLPDGSAPRLCLRLQAFGETLLLEEDSGVQVEGLTVQYLGQAPELLG 120
QY 121 GAEPGYLTGTINGDESVASLHMDGALLGVLYRGAELHQLPLEGTPNSAGGGAHI 180
Db 121 GAEPGYLTGTINGDESVASLHMDGALLGVLYRGAELHQLPLEGTPNSAGGGAHI 180
QY 181 LRRKSPASQCGPWCNKVAPLGSPPRRARRFASLSRFVETLVVADDKMAAFHGGGLKR 240
Db 181 LRRKSPASQCGPWCNKVAPLGSPPRRARRFASLSRFVETLVVADDKMAAFHGGGLKR 240
QY 241 YLLTVAATAAKAFKHPISINPVSIVYTRVYILSGGEGQVGPSPAQOTLRSCAMORGIN 300
Db 241 YLLTVAATAAKAFKHPISINPVSIVYTRVYILSGGEGQVGPSPAQOTLRSCAMORGIN 300
QY 301 TPEDSDPDHFDTAILEFTRODLCGVSTCDTLGNMADVGTVCDDPARSCAIVEDDGLQSAFTAA 360
Db 301 TPEDSGPDHFDTAILEFTRODLCGVSTCDTLGNMADVGTVCDDPARSCAIVEDDGLQSAFTAA 360
QY 361 HELGHVFNMLHDNSKPCISLNGPLSTRHVMAPVMAHVDPEEPWSPCSARFTTDFLDNGY 420
Db 361 HELGHVFNMLHDNSKPCISLNGPLSTRHVMAPVMAHVDPEEPWSPCSARFTTDFLDNGY 420

```

Qy 421 GHCLLDKPEAPLHLPVYGDYKDD 444  
| | | | | | | | | | | | | | | | | |  
Db 421 GHCLLDKPEAPLHLPVYFPCKDYD 444

Job time : 21.7093 secs

## RESULT 15

US-10-174-582-352  
; Sequence 352, Application US/10174582  
; Publication No. US20030027265A1  
; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Collin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3430R1C36  
; CURRENT APPLICATION NUMBER: US/10/174,582  
; CURRENT FILING DATE: 2002-06-18  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 352  
; LENGTH: 837  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-174-582-352

## Query Match

97.3%; Score 2306; DB 9; Length 837;  
Best Local Similarity 98.6%; Pred. No. 5,6e-178;

Matches 438; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MSOTGSHPGKGLAGRWLMGAPCLLPVPLSLVWLLLLSLASPLSARLASPLPREE 60  
| | | | | | | | | | | | | | | | | |  
Db 1 MSOTGSHPGKGLAGRWLMGAPCLLPVPLSLVWLLLLSLASPLSARLASPLPREE 60  
  
Qy 61 IYPERKLNQSVLPQSGTTPARLCLRLQAFETLLLELQDSGVQVEGLTVQYLGQAPELLG 120  
| | | | | | | | | | | | | | | | | |  
Db 61 IYPERKLNQSVLPQSGTTPARLCLRLQAFETLLLELQDSGVQVEGLTVQYLGQAPELLG 120  
  
Qy 121 GAEPGTYLTGTTINGDESVASLHMDGALLGVQYRGAEHLQPLEDGTTPNSAGPGAH 180  
| | | | | | | | | | | | | | | | | |  
Db 121 GAEPGTYLTGTTINGDESVASLHMDGALLGVQYRGAEHLQPLEDGTTPNSAGPGAH 180  
  
Qy 181 LRKSPASGCGPCNKNKAPLGSPSPRRRAKRFASLSRFVETLVVADDKMAAFHAGLKR 240  
| | | | | | | | | | | | | | | | | |  
Db 181 LRKSPASGCGPCNKNKAPLGSPSPRRRAKRFASLSRFVETLVVADDKMAAFHAGLKR 240  
  
Qy 241 YLLTVMAAAAKAKFAPSIRNPVSLVYTRVLIIGSGEGEPQVGPSPAAQTLRSFCAMQGLN 300  
| | | | | | | | | | | | | | | | | |  
Db 241 YLLTVMAAAAKAKFAPSIRNPVSLVYTRVLIIGSGEGEPQVGPSPAAQTLRSFCAMQGLN 300  
  
Qy 301 TPEDSDPHFDFTALFTRODLGCVSTCDTLGMADVTCVCDPARSCAIVEDDGLQSAFTAA 360  
| | | | | | | | | | | | | | | | | |  
Db 301 TPEDSDPHFDFTALFTRODLGCVSTCDTLGMADVTCVCDPARSCAIVEDDGLQSAFTAA 360  
  
Qy 361 HELGHVFNMLHDSKPCISLNGPLSTSRHMAVMAHVDPEPSPCSARFTITDFLDNGY 420  
| | | | | | | | | | | | | | | | | |  
Db 361 HELGHVFNMLHDSKPCISLNGPLSTSRHMAVMAHVDPEPSPCSARFTITDFLDNGY 420  
  
Qy 421 GHCLLDKPEAPLHLPVYGDYKDD 444  
| | | | | | | | | | | | | | | | | |  
Db 421 GHCLLDKPEAPLHLPVYFPCKDYD 444

Search completed: April 21, 2003, 14:55:47















Db	349	AGTAGGAA-----CGAATTATGGGGATNGATCTTCCCGCATCTCGATGTCATCAAC	296
Qy	600	GCGAGGCGCTTACGCTTCGACGAGCCCTGCGCCGCGCGCCAGCTGGCGAAMCCCCGGCTCCA	659
Db	295	GCGAGGCGCTTACGCTTCGACGAGCCCTGCGCCGCGCGCCAGCTGGCGAAMCCCCGGCTCCA	236
Qy	660	CACCGAGGCGCCACGACGATGCTCCGGGGGACACGACACCCGAGGGGAGCGGACGACACTGG	719
Db	235	CATCTGGGCGCCCGAAGAACCCCTCGGTCGACAGTAGATCTTGAAGAGCGCTCAGCGCTGG	176
Qy	720	CCTCGACACTCTCTGGAGCCAGTCCGCTCTCTCCCGCGCGGGGCTCAGAGCCGACGAGCT	779
Db	175	CCCGGACACTCTCTGGAGCCAGTCCGCTCTCTCCCGCGCGGGGCTCAGAGCCGACGAGCT	116
Qy	780	GGTGGGCGCGCGCGCGCTTCATCTCCCGGGCGCCGACGAGTGAAGCTCTTCTGTGG	839
Db	115	GGTGGGCGCGCGCGCGCTTCATCTCCCGGGCGCCGACGAGTGAAGCTCTTCTGTGG	56
Qy	840	CTGACGCGCTCCATGCGCGGCTTGTATGGCCGGGGGCGCTCAGC	881
Db	55	CTGACTGCTTCATGCGCGAGGATGTATGGCGGGGCGCTCAGC	14
RESULT 8			
LOCUS	BB225365		
DEFINITION	BB225365 RIKEN full-length enriched, adult male aorta and vein Mus musculus cDNA A530092A04 3' similar to AF140673 Mus musculus putative secreted metalloprotease AdamTS5 (Adamtss) mRNA, mRNA sequence.	674 bp	mRNA linear EST 31-AUG-2001
ACCESSION	BB225365		
VERSION	BB225365.2	GI:15410179	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 674)		
AUTHORS	Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Himoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Koude,O., Koyama,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasak,D., Shibata,K., Shinagawa,A., Shiraki,Y., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M., and Hayashizaki,Y.		
TITLE	RIKEN Mouse ESTs (Arakawa,T., et al. 2001)		
JOURNAL	unpublished (2001)		
COMMENT	On Jul 1, 2000 this sequence version replaced gi:889376. Contact: Yoshinide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/ Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y. Normalization and subtraction of cap-trapper selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) Wegli,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanabe,K., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A., and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y., and Hayashizaki,Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a		

	<p>nonredundant cDNA library Genom. Res., 11 (2), 281-289 (2001) Yamanaka, I., Kiyosawa, H., Kondo, S., Saito, T., Shinagawa, A., Aizawa, 'K.', Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., Arakawa, T., Ishii, Y. and Hayashizaki, Y. Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct. Funct. Genomics 2 pre, L72-L86 (2001) Please visit our web site (<a href="http://genome.gsc.riken.go.jp/">http://genome.gsc.riken.go.jp/</a>) for further details.</p>					
	<p>cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.</p>					
FEATURES						
source	Location/Qualifiers					
	1..674					
	/organism="Mus musculus"					
	/db_xref="taxon:10090"					
	/clone="A530092A04"					
	/clone_lib="RIKEN full-length enriched, adult male aorta and vein"					
	/sex="male"					
	/tissue_type="aorta and vein"					
	/dev_stage="adult"					
	/lab_host="DH10B"					
	/note="Site_1: SalI; Site_2: BamHI: cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGGAGAGAAGATCCAAAGACCTTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTTCGTCAATAATTAATATCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified Bluescript KS(+) after bulk excision from Lambda FLC 1."					
BASE COUNT	160 a	174 c	183 g	157 t		
ORIGIN						
Query Match	21.1%; Score 320.2; DB 10; Length 674;					
Best Local Similarity	86.1%; Pred. No. 1.9e-58;					
Matches 378; Conservative	0; Mismatches 58; Indels 3; Gaps 2					
OY	1023	CCACACTGAAGAAGCTTTGGCAAGTGCGACGACCACAACACACAGCCTGGAGATGACCATG	1082			
Db	19	CCCGCACCTCAAAAACCTTTTGCMAAATGACGCCCAACATPACCGTAGG--GGATATCCCG	76			
OY	1083	AGGACCACTAGCATGCAGCTATCCGTGTTACTTCGCGGAGGATTTATGTGGCANTCATCAT	1142			
Db	77	AAGACCCCTACCATGACGACATCCGTGTCGCCGAGAGGATTTATGTGGCATCATCAT	136			
OY	1143	GTCGACACCTGGGATGCGAGACGCTTGGGACCATPATTGTTCTCCAGAGCCACACTGTGCTG	1202			
Db	137	GTGAC -CCCTGGGATGCGAGACGCTTGGGACCATPATGTTCTCCGAGGCCACACTGTGCTG	195			
OY	1203	TGATTGAAGAGAGATGGCTCCAGCGAGCCCTTGCTGCTGCTGACGAAATGAGACATTAC	1262			
Db	196	TGATTGAAGATGATGGCTCCATGACAGCCTTCACGTGTGCTCATGAAATTTGGCATCTCAC	255			
OY	1263	TTGGCCCTCCCATACGATTTCCAATTCGTGGAAGAGACCTTTGGTTCACAGAAAGATA	1322			
Db	256	TTGGCCCTTCTCATAGCAATTCMAATTCGTGAAGAGAACTTGGTATCAAGAGACA	315			
OY	1323	AGCGCTTATATGCTCCATCCATCCTTACGACATTTGATGATCATTAAGCCCTGGTCCAATGCA	1382			
Db	316	AGCGTTTATATGCTTTCATTCCTTTACGACATGATGATCATCAAAGCCCTGGTCCAATGCA	375			
OY	1383	CTTACGCCACCATCAAGAAATTCGTGATGATGAGGACCATGTAACCTGTTGGCTGGAAGCTTAC	1442			

```

Db      376   COTCAGCCACCATTCACAGATAATCTCGATGATGGTGCATGTGTAATTGTTGGTAGACCTAC 435
Oy      1443  CACGAAGAACAAGATCTCTGGG 1461
Dn      436   CACGAAGAACAAGATTTGGG 454

RESULT 9
BE863546/c
LOCUS
DEFINITION BE863546 .467 bp mRNA linear EST 29-SEP-2000
UT-M-BH0-ak-h-c-12-0-U-T1 NIH_BMAP_M_ST1 Mus musculus cDNA clone
UT-M-BH0-ak-h-c-12-0-U1 5', mRNA sequence.
Accession BE863546
Version BE863546
Keywords BE863546.1 GI:10383632
Source EST.
Organism house mouse.
Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 467)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 9704447
COMMENT Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-1790, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mestr@nimh.nih.gov
cDNA Library Preparation: M.B. Soares Lab Clone distribution:
Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
should be noted that Bento Soares is generating a small number of
additional specialized non-redundant arrays of BMAP cDNAs whose
availability will be considered under appropriate and limited
collaborative arrangements
Seq primer: M13 Reverse.
Location/Qualifiers
I..467
FEATURES
source
```

```

/db_xref="taxon:10090"
/clone_1ib="M-BH0-akh-c-12-0-U1"
/clone_1ib="N1H_BMAP_M.S1"
/seq_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: pRT3D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; The
N1H_BMAP_M.S1 library is a subtracted library derived from
a mixture of normalized libraries from ten regions of the
mouse brain (cerebellum, brain stems, olfactory bulbs,
hypothalamus, cortex, amygdala, basal ganglia, pineal
gland, striatum, hippocampus). The driver used for
subtraction consisted of a pool of 20,000 cDNA clones
obtained from non-normalized and normalized libraries of
these ten regions of the mouse brain."

BASE COUNT      99 a      143 c      139 g      86 t
ORIGIN

```

Oy	441	CTGTGGCTGCTTTTACCTCTGTGGGGTCTCGACGGCTTCTTCCGCTCAACAGCGGC	500
Db	347	CCCTAGCTGTCTTTTACCTCTGTGGGGGTCTCGATGGCTTCTTTCACATCAAGCATGTGGC	288
Oy	501	GCTACACCTTAAAGCAGCTGCTGTCCCGGACCTGTGGGCGAGAAAGAAAGGGGCCCTGT	560
Db	287	GCTACACTCTAAGCAGCTCCTGTGCTGGGTCTTGGGCGAGAGTATAA-----CGAATTT	234
Oy	561	ACGGGGATGGGTCCGCACGAGATCTGCACGTTCTACACCCGGGAGGCTTCAGCTTCGAGG	620
Db	233	ATGGGAGTGGATCTTCCGGCATCTCGCATGTCTACAAACGGGAGGCTTTAGCTTCGAGG	174
Oy	621	CGCTCCCGCGCGGCCACAGCTGCGAAACCCCGCTCCACACCGGAGGCCACAGAGCATG	680
Db	173	CCCTCCCGCCACGGCCCATGTTGCCGAGCTCTCTGCATCCCATCTTGGGCCCGAGGAGCC	114
Oy	681	CTCCGGGCGACAGCACCCGAGCGGAGCGGACGACCTGGCTCCGAGCTCTTTGGACCAT	740
Db	113	CTTCGGTGCACAGTAGATCTTAGGAGACCTCAAGCCCTGAGCCCCCGAGCTGCTGGACCACT	54
Oy	741	CCGCTCTCTCGCCCTGTGGGGCTAGAGACCGCAACAGCTGGTGGCGCGCGG	793
Db	53	CAGCTTCTCCGCACTCTGGGAACGGGGAGCTCAGACACTTTGGTGGAGCGCTAGG	1

RESULT	10				
LOCUS	AL552429	997 bp	mRNA	linear	EST 16-FEB-2001
DEFINITION	AL552429 LTI_NF1006_P12 Homo sapiens cDNA clone CS001070YJ19 5 prime, mRNA sequence.				
ACCESSION	AL552429				
VERSION	AL552429.1	GI:12891319			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
AUTHORS	Li, W.B., Gruber, C., Jesssee, J. and Polayres, D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished (2001)				
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seque@genoscope.cns.fr, web : www.genoscope.cns.fr.				
FEATURES	Location/Qualifiers				
source	1..997				

/clone\_11b="L1\_NFL006\_PL2"  
 /tissue="cardiovascular"  
 /tissue\_type="placenta"  
 /note="vector: PCMVSPORT 6; Site\_1: NotI; 1st strand cDNA  
 was primed with a NotI-oligo(NT) primer. Five prime end  
 enriched, double-stranded cDNA was digested with Not I and  
 cloned into the Not I and Eco RV sites of the PCMVSPORT 6  
 vector. Library was normalized. Library was constructed by  
 Life Technologies. Contact : Feng Liang Life Technologies,  
 a division of Invitrogen 9800 Medical Center Drive  
 Rockville, Maryland 20850, USA Fax : (1) 301 610 8371  
 Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com

BASE COUNT	175 a	337 c	274 g	205 t	6 others
ORIGIN					
Query Match	17.5%	Score 265.2	DB 9	Length 997	
Best Local Similarity	95.6%	Pred. No. 1.3e-46			
Matches 323	Conservative 1	Mismatches 9	Indels 5	Gaps 5	

Query Match 17.5%; Score 265.2; DB 9; Length 997;  
 Best Local Similarity 95.6%; Pred. No. 1,36-46;  
 Matches 325; Conservative 1; Mismatches 9; Indels 5; Gaps 5;









## ORGANISM

Mus musculus

REFERENCE  
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 337)  
Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,  
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person  
, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Rifter  
, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,  
Waterston, R. and Wilson, R.  
The WashU-NCI Mouse EST Project 1999

TITLE  
JOURNAL

Unpublished (1999)

## COMMENT

Contact: Marra M/WashU-NCI Mouse EST Project 1999  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
This clone was previously sequenced on the 5' end only, this new  
data is from the 3' end  
Possible reversed clone: similarity on wrong strand  
High quality sequence stop: 304.

## FEATURES

## source

Location/Qualifiers  
1..337  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:569515"  
/clone\_lib="Life Tech mouse embryo 8 5dpc 10664019"  
/tissue\_type="embryo"  
/dev\_stage="8.5dpc embryos"  
/lab\_host="DH10B"  
/note="Organ: whole embryo; Vector: PCMV-SPORT2; Site\_1:  
Salt: Site\_2: Not; Cloned unidirectionally. Primer:  
Oligo dT: 8.5dpc embryos. PCMV-SPORT2 vector."  
BASE COUNT 61 a 95 c 110 g 71 t  
ORIGIN

## Query Match

Best Local Similarity 80.1%; Score 215.2; DB 9; Length 337;  
Matches 266; Conservative 0; Mismatches 63; Indels 3; Gaps 1;

OY 191 GCCCAGCCGCCACCCCTGGCGCGCGGCGCAGAGCAAGGGCTGTCAGCAACATCGAC 250  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 5 GCCCATTTACAAACCTTGCGCGCGGCGAGCGCAGAGCGGGGCTGTCCAGAAATATAGAC 64  
OY 251 CAACTCTACTCGCGCGCGCAAGTGGCTACTGCTCTACGCGGGCGCGCGCAGGTTTC 310  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 65 CAACTCTACTCTGGCGGTGGCAAGTGGGCTACTGTTCTACGCGGGCGCGCGAGGTTTC 124  
OY - 311 CCTTGGACCTGGAGAGATGGTGGGTCATTGGTG---GCTTGGTGGCGCGCAGGA 367  
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 125 CTGCTGAGACTGAGAGATGATACAGAGGGTGGTGGTGTGAGCATGTTACTGCAGGA 184  
OY 368 GCGGGGAGAGTGGCCCTGGCGCGCGGCGCAGACACTGTTCTATCGGGCGCAGAGTGGAC 427  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 185 GGAGGGCTGAGGCGATCCCTGCGCACCGGGGTCACTGTTCTACAGAGCGACCGTGGAC 244  
OY 428 GGTAGTCCCGCTCTGTGGCTGCTTTGACTCTGTGGGGGTCTGACAGCGCTTTCCGG 487  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 245 GGCACCCCTCGATCCCTAGCTGTCTTTGACCTCTCGGGGGTCTGATGGCTTTTGA 304  
OY 488 GTCAAGCAAGCGCGGTACACCTTAAGCCACT 519  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 305 GTCAAGCAAGCGCGGTACACTTAAGCCACT 336

Search completed: April 26, 2003, 04:36:07  
Job time : 4653.37 secs



```

188 GGAGCA-----GCCCCATTACACCTTGGCCGGGACGAGGCGGGCG 235
232 GGTGGTGAACATGACCACTACTCCGGGGGGGAGTGGGTACTGGTCTA 291
236 GCTGGTCAATATAGACCAACTACTGGGGGGGAGGAGTGGGTACTGGTCTA 295
292 CGCGGGGGGAGGTTCTCTTGGACCTGGAGAGATGGTTCGGGGGATTCGTC- 350
296 CGCGGGGGGAGGTTCTCTTGGACCTGGAGAGATGACAGTGGGTGGTCTCTG 355
351 --GTTTCCTGGGGGAGGAGGAGGAGTGGGGGCGCCGACCGGACGACCTCTT 408
356 TAGCATCTTACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 415
409 CTATCGGGGACAGTGGAGGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAG 468
416 CTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 475
469 TCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 528
476 TCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 535
529 ACCCTGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 588
536 GTCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 589
589 CCTGTACACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 648
590 TGTGTACACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 649
649 CCGCGGCTTCACACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 708
650 TCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 709
709 CGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 768
710 CTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 769
769 ACCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 828
770 ACTGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 829
829 GCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 888
830 CCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 889
889 GGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 948
890 GGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 949
949 CCGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1008
950 CGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1009
1009 CAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1068
1010 CAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1069
1069 GGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1128
1070 AGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1129
1129 TGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1188
1130 TGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1189
1189 GCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1248
1190 GCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1249
1249 AATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1308

```

```

Db 1250 AATTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1309
Qy 1309 TTCCACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1368
Db 1310 TACTACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1369
Qy 1369 CTGGTCAATATGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1428
Db 1370 CTGGTCAATATGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1429
Qy 1429 TTTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1461
Db 1430 TTTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1462

```

## RESULT 2

```

us-10-105-929-1
Sequence 1, Application US/10105929
Patent No. US20020137142A1
GENERAL INFORMATION:
APPLICANT: Holzman, Douglas A.
APPLICANT: Goodheart, Andrew D.J.
TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
FILE REFERENCE: 09404/041001
CURRENT APPLICATION NUMBER: US/10/105,929
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/130,491
PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-07
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/058,108
PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-05
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/054,961
PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-06
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 4676
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (460)...(3360)
us-10-105-929-1

```

```

Query Match 17.6%; Score 267; DB 12; Length 4676;
Best Local Similarity 53.4%; Pred. No. 1.1e-65;
Matches 586; Conservative 0; Mismatches 505; Indels 6; Gaps 1;

```

```

Qy 363 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 422
Db 785 CCGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 844
Qy 423 TGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 482
Db 845 TGAATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 904
Qy 483 TGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 542
Db 905 TCTTACTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 964
Qy 543 AAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 602
Db 965 CCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1024
Qy 603 AGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 662
Db 1025 ATCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1084
Qy 663 CGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 722
Db 1085 GGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1144
Qy 723 CGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 782

```

Db 1145 GGTGCGCCGAGGACCCGCGACATGCAAGGCGGTAGAGACGCCACAGGA-----ACTGGAA 1198  
QY 783 GCGCGCGCGCGCGCGCTCCATCTCCCGGGCCCGCAGGTGGAGTGGCTTCTGTGGCTG 842  
Db 1199 GCATAGAAAGAAAGCATTTGTGTCTCAGTCACCGCGCATGTGGAACCATGCTTGTGGCAG 1258  
QY 843 ACGCTCCATGGCGCGGTGTATGCGCGGGGCGCTGACGATTACCTGCTGACCTGGCCT 902  
Db 1259 ACCAGTCTGATGCGAATATCCACGGCAGTGTCTAAAGCATTAACCTCTCAGCTGTCTTT 1318  
QY 903 CCATGGCCCAATAGGCTGTACAGCATGCTAGCATGAGAACCATCCCGCTGGCGCTGG 962  
Db 1319 CGGTGGCAGCCGAGTGTGCAAAACACCCGACATGCTGTAACTTACGTTACGCTGTGGTGG 1378  
QY 963 TGAAGGTGTGTGTGTAGCGACAAAGACAGACGCTGGAGTGGACCAAGACGCTGCCA 1022  
Db 1379 TGAAGATCTTGGTTCATCCACGATGAAACAGAAAGGCGCGGAAGTACCTCCATGCTGCC 1438  
QY 1023 CCACACTCAAGACTTTTTCGAAGTGGCAGACCAACACAGCAGCTGGAGATGACCATG 1082  
Db 1439 TCACCTGTGGGAACCTTTTTCACACTGGCAGACAGACCAACCCAGCTGACCGGGATG 1498  
QY 1083 AAGAGACTAGCATGAGCTATCTGTGTACCTGCGAGGATTTATGTGGCATCATTCAT 1142  
Db 1499 CAGAGCACTATGACACAGCAATCTTTTCCAGACAGACGACTTGTGGTCCAGACAT 1558  
QY 1143 GTGACACCTGGGAATGGCAGACGTTGGAGCATATGTTCTCCAGAGCGCAGCTGTGTG 1202  
Db 1559 GGTATCTTGTGGATGGCTGATGTTGGAACTGTGTATCCGACAGAAAGCTGCTCGG 1618  
QY 1203 TGAATGAGACATGAGCTTCCACGACGCTTCACTGTGCTCAGCAAAATGGAATTTAC 1262  
Db 1619 TCATGAAATGATGATGTTTACAGGCTGCTTCCACAGACCCATGAAATAGGCGACGTG 1678  
QY 1263 TGGGCTCTCCCATGCAATTCGTGTGAAGAGACCTTTGGTTCACAGAAAGATA 1322  
Db 1679 TTAACATGCCACATGATGTGTGAAGACAGTGTGCCAGCTTATGTGTGAACCAAGATT 1738  
QY 1323 ACGGCTTAATGCTTCATCCTTACAGCATGTATGATCATTAAGCCGCTGGTCAATGCA 1382  
Db 1739 CCCACATGATGGCGTCAATGCTTTCACCTGAGCACAGCCGCTTGTGCTCCCTGGA 1798  
QY 1383 CTTCAGCCCACTCAGCAATTCCTGTGATGATGGCCATGTATGTTTCTGTGGACCTAC 1442  
Db 1799 GTCCTACATGATGATCATCTTCTGTGAATATGTCATGGGAGATGTTTGTGAGACAAC 1858  
QY 1443 CAGGAAAGCAGATCTG 1459  
Db 1859 CTCAGATCCCATACAG 1875

RESULT 3  
US-10-163-316-3  
; Sequence 3, Application US/10163316  
; Publication No. US20020197703A1  
; GENERAL INFORMATION:  
; APPLICANT: Kapeller-Libermann, Rosana  
; TITLE OF INVENTION: 65552, A Human Matrix Metalloproteinase and Uses  
; FILE REFERENCE: MP101-025P1RMN  
; CURRENT APPLICATION NUMBER: US/10/163,316  
; PRIOR FILING DATE: 2002-06-05  
; PRIOR APPLICATION NUMBER: 60/297,863  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 2469  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(2469)

US-10-163-316-3  
Query Match 17.0%; Score 257.8; DB 9; Length 2469;  
Best Local Similarity 61.3%; Pred No. 3.7e-63;  
Matches 415; Conservative 0; Mismatches 262; Indels 0; Gaps 0;

QY 780 GGTGGCGCGCGCGCGCCCTCCATCTCCCGGGCCCGCAGGTGGAGTGGCTTCTGTGGTGG 839  
Db 617 GGTGTGGCGCGCGCGCAAGCGTTTGTGTGTATCCCGGGGTAGTGGAGAGCGTGTGGTGG 676  
QY 840 CTGACCGCTTCATGGCGCGCTGTATGCGCGGGGCTGACGATTAATCTGTGACCTGG 899  
Db 677 CGAGCAGCTCAATGCTCAATTTCCACGGCGGACCTGGAAATATCTGTGACGCTCC 736  
QY 900 CCTCCATGCCAATAGGCTGTACAGCATGCTAGCATGAGAACCAATCCGCTGGCGG 959  
Db 737 TGGCAAGCGCGCGCGCTCTATCCGATCCCAAGCTTCAACCCCATCAACATCTTGG 796  
QY 960 TGTGTAAGTGTGTGTGTAGCGACAGACAGAGAGCCTGGAGTGGACCAAGACGCTG 1019  
Db 797 TGTGTAAGTGTGTGTGTGTAGATGATGATGATGATGATGATGATGATGATGATGATG 856  
QY 1020 CCACCACTCAAGAACTTTTTCAGAGTGGCAGACCAACACAGCTGGAGATGACC 1079  
Db 857 CCTGACGCTGGGCAACTCTGTGCTGCTGAGAGAGCTGAACAAATGAGTGAACAGC 916  
QY 1080 ATGAGAGGACTAGCATGAGCTATCTGTGTACTGCGGAGGATTTATGTGGCATCAT 1139  
Db 917 ACCCGAGTACTGTGACACTGCCATCTCTTCCACAGGACGACCTGTGTGGACCA 976  
QY 1140 CATGTGACACCCCTGGGAATGGCAGAGCTTGGGACATATGTTCTCCAGAGCGCAGCTGTG 1199  
Db 977 CCGTGTACACCTGGGCGATGCGTGTATGATGATGATGATGATGATGATGATGATGATG 1036  
QY 1200 CTGTGATTAAGACATGAGCTTCCACGACGCTTCACTGTGCTCAGAAATGGAATTTAC 1259  
Db 1037 CTGTGATTAAGACATGAGCTTCCACGACGCTTCACTGTGCTCAGAAATGGAATTTAC 1096  
QY 1260 TACTTGGCTCTCCCATGAGATTCCTCAATTCGTGAAGAGACCTTGTGGTCCACAGAG 1319  
Db 1097 TGTTCACATGATGCGCGCATGATGATGATGATGATGATGATGATGATGATGATGATG 1156  
QY 1320 ATGAGGCTTAATGCTTCATCCTTACAGCATGTATGATCATTAAGCCCTGGTCCAAAT 1379  
Db 1157 CCACCACTGATGTGCCGACCTCATCATCATCATCATCATCATCATCATCATCATCAT 1216  
QY 1380 GCACCTCAGCCACCATCAGAAATTCCTGTGATGATGATGATGATGATGATGATGATGATG 1439  
Db 1217 GCAGTGTGCTCATCATCAGCATCTTCTGTGACAGCGGGACGATGATGATGATGATGATG 1276  
QY 1440 TACCAAGAGCAGATC 1456  
Db 1277 AACCCAGCAAGCCCATC 1293

RESULT 4  
US-09-965-631-3  
; Sequence 3, Application US/09965631  
; Patent No. US20020115842A1  
; GENERAL INFORMATION:  
; APPLICANT: Friedle, Carl Johan  
; APPLICANT: Hilbun, Erin  
; TITLE OF INVENTION: No. US20020115842A1 Human Proteases and Polynucleotides Enco  
; FILE REFERENCE: Lex-0241-USA  
; CURRENT APPLICATION NUMBER: US/09/965,631  
; PRIOR FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: US 60/236,689  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 2853  
; TYPE: DNA





;; CURRENT APPLICATION NUMBER: US/09/965,631  
 ;; CURRENT FILING DATE: 2001-09-27  
 ;; PRIOR APPLICATION NUMBER: US 60/236,689  
 ;; PRIOR FILING DATE: 2000-09-29  
 ;; NUMBER OF SEQ ID NOS: 7  
 ;; SOFTWARE: FastSeq for Windows Version 4.0  
 ;; SEQ ID NO: 7  
 ;; LENGTH: 3446  
 ;; TYPE: DNA  
 ;; ORGANISM: homo sapiens  
 US-09-965-631-7

Query Match 17.0%; Score 257.8; DB 10; Length 3446;  
 Best Local Similarity 61.3%; Pred. No. 4,1e-63;  
 Matches 415; Conservative 0; Mismatches 262; Indels 0; Gaps 0;

Qy 780 GGTGGCGGGCGGGCGGCTTCATCTCCGGGCGCCGAGGTGAGTGTCTCTGGTG 839  
 Db 1013 GGTCTGGCGCGCCAGAGGTTTCTGTCTATCCCGGTAGCTGAGACGCTGTGTCG 1072  
 Qy 840 CTGACGCGCTCATGCGCGGCTTGTATGCGCGGCGCTGACAGATTAAGTCTGACCTGG 899  
 Db 1073 CGGACGATCATGTGTCAAGTTCACGGCGCGGAGCTGGAACATTAATCTGCTGACCTGC 1132  
 Qy 900 CCTCATCGCAATAGCTGTACAGCATGCTAGCATGAGAACCATCCGCGCTGGCGG 959  
 Db 1133 TGGCAAGCGGCGCGGCTTACCGCATCCAGCATCTCTCAACCCATCAATCGTTG 1192  
 Qy 960 TGGTGAAGTGTGTGTAGCGGCAAGACAGAGAGCTGGAATGAGCAAGACGCTG 1019  
 Db 1193 TGGTGAAGTGTGTGTGTAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1252  
 Qy 1020 CCACGACACTCAAGAACTTTTGAAGTGGAGACAGACACACACAGCAGTGGAGATGAC 1079  
 Db 1253 CCGTGAAGCTGCGCAACTTGT 1312  
 Qy 1080 ATGAGGACACTAGCATGACATCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1139  
 Db 1313 ACCCGGACTAGTGGGACACTGTGCATCTCTTCCAGCAGGACCTGTGTGTGTGTGTGTGT 1372  
 Qy 1140 CATGAGACCTGGAGATGGAGAGCTTGGAGCATATGTTCTCCAGAGCGGACCTGTG 1199  
 Db 1373 CCTGAGACACCTGGGAGATGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1432  
 Qy 1200 CTGTGATTTGAAGACATGAGCTCCACGAGCCTTCACTGTGCTGCTGCAAGAAATCGGACAT 1259  
 Db 1433 CTGTGATTTGAAGACATGAGCTCCACGAGCTTCACTGTGCTGCTGCAAGAAATCGGACAT 1492  
 Qy 1260 TACTTGGCTCTCCATGACATGATTCGAAATTCGTGTGAAGAGACCTTGTGTGTGTGTGTGT 1319  
 Db 1493 TGTTCACATGAGCTCCATGACATGTGAAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1552  
 Qy 1320 ATAAAGGTTAATGTTCTTCATCTTACAGATGATGATCTTAAGCGCTGTGTGTGTGTGTGT 1379  
 Db 1553 CCACACCATGATGTTCTCCAGACCTTCATGATGATGATGATGATGATGATGATGATGATGAT 1612  
 Qy 1380 GCACTTCAAGCAGCATCAAGAAATTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1439  
 Db 1613 GCAAGT 1672  
 Qy 1440 TACGACGAAGACGATC 1456  
 Db 1673 AACCCAGCAAGCCATC 1689

RESULT 7  
 US-10-097-597-13  
 ; Sequence 13, Application US/10097597  
 ; Publication No. US2003002352a1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hirose, Kunitaka  
 ; APPLICANT: Inoguchi, Elji  
 ; APPLICANT: Hakozaaki, Michinori

;; APPLICANT: Ishioka, Keiko  
 ;; APPLICANT: Ishida, Yukako  
 ;; APPLICANT: Matsushita, Kouji  
 ;; APPLICANT: Kuno, Kouji  
 ;; TITLE OF INVENTION: human ADAMTS-1 protein, gene encoding the same,  
 ;; TITLE OF INVENTION: pharmaceutical  
 ;; FILE REFERENCE: 057092  
 ;; CURRENT APPLICATION NUMBER: US/10/097,597  
 ;; PRIOR APPLICATION NUMBER: 09/445,023  
 ;; PRIOR FILING DATE: 1999-12-03  
 ;; PRIOR APPLICATION NUMBER: JP 9-160422  
 ;; PRIOR FILING DATE: 1997-06-03  
 ;; NUMBER OF SEQ ID NOS: 14  
 ;; SOFTWARE: PatentIn version 3.0  
 ;; SEQ ID NO: 13  
 ;; LENGTH: 2184  
 ;; TYPE: DNA  
 ;; ORGANISM: Mus sp.  
 ;; FEATURE:  
 ;; NAME/KEY: exon  
 ;; LOCATION: (1)..(2184)  
 US-10-097-597-13

Query Match 16.9%; Score 256.8; DB 9; Length 2184;  
 Best Local Similarity 61.8%; Pred. No. 6,9e-63;  
 Matches 408; Conservative 0; Mismatches 252; Indels 0; Gaps 0;

Qy 797 GCGTCATCTCCGCGGCGCGGAGTGAAGTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 856  
 Db 34 GCAATTTGT 93  
 Qy 857 CGGT 916  
 Db 94 GACTTCCAGGCGGCGGCTTAAAGCATTTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 153  
 Qy 917 CTGTGAGCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 976  
 Db 154 TTTTCAAGCATCCAGCATTTAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 213  
 Qy 977 CTAGGCGACAGACAGACCTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1036  
 Db 214 ATATACGAGGAGCAGAGGAGCAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 273  
 Qy 1037 TTTTCAAGT 1096  
 Db 274 TTTTCAAGT 333  
 Qy 1097 GCAGTATCTGT 1156  
 Db 334 ACTGCAATCTGT 393  
 Qy 1157 ATGCGAGCTGT 1216  
 Db 394 ATGCGAGATGT 453  
 Qy 1217 GCGCTCCAGGCGCTTCACTGT 1276  
 Db 454 GGTGTGCAAGCTGT 513  
 Qy 1277 GAGCATTTCAATTTGTGTGAAGAGACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1336  
 Db 514 GATGATGTGAAGACACTGT 573  
 Qy 1337 TCCATCTTACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1396  
 Db 574 TCGATGCTCTCCAGCTTAAAGCATGAGCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 633  
 Qy 1397 ACAGAAATTCGT 1456  
 Db 634 ACCTCTTCTGT 693

## RESULT 8

US-10-097-580-13  
: Sequence 13, Application US/10097580  
: Publication No. US20030032168A1

## : GENERAL INFORMATION:

: APPLICANT: Hirose, Kunitaka  
: APPLICANT: Inoguchi, Eiji  
: APPLICANT: Hakozeaki, Michinori  
: APPLICANT: Ishioke, Keiko  
: APPLICANT: Ishida, Yukako  
: APPLICANT: Matsushima, Kouji  
: APPLICANT: Kuno, Kouji

: TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical  
: FILE REFERENCE: 057092  
: CURRENT APPLICATION NUMBER: US/10/097,580

: PRIORITY FILING DATE: 2002-03-15  
: PRIOR APPLICATION NUMBER: 09/445,023  
: PRIOR FILING DATE: 1999-12-03  
: PRIOR APPLICATION NUMBER: JP 9-160422

: NUMBER OF SEQ ID NOS: 14  
: SOFTWARE: PatentIn version 3.0  
: SEQ ID NO: 13  
: LENGTH: 2184  
: TYPE: DNA  
: ORGANISM: Mus sp.

: FEATURE:  
: NAME/KEY: exon  
: LOCATION: (1)..(2184)  
US-10-097-580-13

Query Match 16.9% Score 256.8; DB 9; Length 2184;  
Best Local Similarity 61.8%; Pred. No. 6.9e-63;

Matches 408; Conservative 0; Mismatches 252; Indels 0; Gaps 0;

```
OY 797 CGCTTCATCTCCGGGCGCCGAGGTGAGCTGCTTGTGTGGTGGACGCGCTCATGGCG 856
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 34 CGATTGTGTCCAGCCCGCTTATGTGMAACCATGCTGCTGTAAGACACATGTCATGGCC 93
OY 857 CGCTTGTATGCGGGGCGCTGACGATTAACCTGCTGACCTGGCTCCATCCCAATAG 916
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 94 GACTTCCACGGGAGGCGTCTAAGCATTAACCTTACCCCTGTTCCGTTGGCAGCCAGG 153
OY 917 CTGTACGACCATGCTAGTATGAGAACCAATCCGCTGCGCGGTGTAAGTGTGGTG 976
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 154 TTTTCAAGCATCCAGCATTTAGGAATTCATTAAGCTGTGTGTGTGTGAAGATCTTGGTC 213
OY 977 CTAGCGACAAAGACAGAGCGCTGAGTGAAGAGAGCGCTGCCACCACTCAAGAAC 1036
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 214 ATATACGAGAGACAGAGAGGACCAAGATTAACCTTACATGACAGCTTCACCTCGAAT 273
OY 1037 TTTTCAAGTGGCAGACCAACACAAAGCTGGAGATGACATAGAGAGCACTACAGAT 1096
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 214 TTTTCAAGTGGCAGACCAACACAAAGCTGGAGATGACATAGAGAGCACTACAGAT 333
OY 1097 GCAGTATCTCTTACTCGGAGGATTTATGTGGCATCTTATTCATGACACCTGGGA 1156
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 334 ACTGCAATTCCTTACCAAGACAGATTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 393
OY 1137 ATGCGAGAGTGGAGCATATGTTCTCAGAGCGGAGCTGTGTGTGTGTGTGTGTGTGT 1216
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 394 ATGCGAGAGTGGAGCATATGTTCTCAGAGCGGAGCTGTGTGTGTGTGTGTGTGTGT 453
OY 1217 GGCCTCCACGACGCTTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 454 GGTTCGAGCTGCTTACCAAGACAGATTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 513
OY 1277 GACGATTCGAATTCCTGTGAAGAGACCTTTGGTTCACAGAGATTAAGCGGTATGTCT 1336
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 514 GATGATGTAAGCACTGTGCGCAGCTTGAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 573
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

## RESULT 9

US-09-445-023A-13  
: Sequence 13, Application US/09445023A  
: Patent No. US20020119167A1

## : GENERAL INFORMATION:

: APPLICANT: Hirose, Kunitaka  
: APPLICANT: Inoguchi, Eiji  
: APPLICANT: Hakozeaki, Michinori  
: APPLICANT: Ishioke, Keiko  
: APPLICANT: Ishida, Yukako  
: APPLICANT: Matsushima, Kouji  
: APPLICANT: Kuno, Kouji

: TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical  
: FILE REFERENCE: 057092  
: CURRENT APPLICATION NUMBER: US/09/445,023A  
: PRIORITY FILING DATE: 1999-12-03  
: PRIOR APPLICATION NUMBER: JP 9-160422

: NUMBER OF SEQ ID NOS: 14  
: SOFTWARE: PatentIn version 3.0  
: SEQ ID NO: 13  
: LENGTH: 2184  
: TYPE: DNA  
: ORGANISM: Mus sp.

: FEATURE:  
: NAME/KEY: exon  
: LOCATION: (1)..(2184)  
US-09-445-023A-13

Query Match 16.9% Score 256.8; DB 10; Length 2184;  
Best Local Similarity 61.8%; Pred. No. 6.9e-63;

Matches 408; Conservative 0; Mismatches 252; Indels 0; Gaps 0;

```
OY 797 CGCTTCATCTCCGGGCGCCGAGGTGAGCTGCTTGTGTGGTGGACGCGCTCATGGCG 856
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 34 CGATTGTGTCCAGCCCGCTTATGTGMAACCATGCTGCTGTAAGACACATGTCATGGCC 93
OY 857 CGCTTGTATGCGGGGCGCTGACGATTAACCTGCTGACCTGGCTCCATCCCAATAG 916
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 94 GACTTCCACGGGAGGCGTCTAAGCATTAACCTTACCCCTGTTCCGTTGGCAGCCAGG 153
OY 917 CTGTACGACCATGCTAGTATGAGAACCAATCCGCTGCGCGGTGTAAGTGTGGTG 976
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 154 TTTTCAAGCATCCAGCATTTAGGAATTCATTAAGCTGTGTGTGTGTGTGTGTGTGTGT 213
OY 977 CTAGCGACAAAGACAGAGCGCTGAGTGAAGAGAGCGCTGCCACCACTCAAGAAC 1036
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 214 ATATACGAGAGACAGAGAGGACCAAGATTAACCTTACATGACAGCTTCACCTCGAAT 273
OY 1037 TTTTCAAGTGGCAGACCAACACAAAGCTGGAGATGACATAGAGAGCACTACAGAT 1096
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 214 TTTTCAAGTGGCAGACCAACACAAAGCTGGAGATGACATAGAGAGCACTACAGAT 333
OY 1097 GCAGTATCTCTTACTCGGAGGATTTATGTGGCATCTTATTCATGACACCTGGGA 1156
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 334 ACTGCAATTCCTTACCAAGACAGATTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 393
OY 1137 ATGCGAGAGTGGAGCATATGTTCTCAGAGCGGAGCTGTGTGTGTGTGTGTGTGTGT 1216
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 394 ATGCGAGAGTGGAGCATATGTTCTCAGAGCGGAGCTGTGTGTGTGTGTGTGTGTGT 453
OY 1217 GGCCTCCACGACGCTTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db 454 GGTTCAGAGCTGCTTCACACACACCCCATGATTGGCCATGTTTACATGCCGCAC 513  
OY 1277 GACGATTCGAATTCGTGAAGAGACCTTGGTTCACAGAAAGATAGCGCTTATGCT 1336  
Db 514 GATGATGCTGAAGCACTGTGCCAGCTGTGAATGTGTGAGTGGCGATTCATCTGATGTC 573  
OY 1337 TCCATCTCTACCAATGATGATGATCTTAAGCCCTGGTCCAAATGACATTCAGCCATC 1396  
Db 574 TCGATGCTCTCAGCTTGAACCATAGCCAGCCCTGGTCCACTTGGAGTGGCTACATGCTC 633  
OY 1397 ACAGATTCCTGAGATGAGCGGCACTGTTGCTGAGCACTTACCCAGAAAGAGATC 1456  
Db 634 AGTCCTCTAGATTAAGACAGCGGGAAATCTTGAATGACAAAGCCCAAGATCCATC 693

## RESULT 10

US-10-097-597-2

Sequence 2, Application US/10097597  
Publication No. US20030022352A1  
GENERAL INFORMATION:  
APPLICANT: Hirose, Kunitaka  
APPLICANT: Inoguchi, Eiji  
APPLICANT: Hakozaeki, Michinori  
APPLICANT: Ishioke, Keiko  
APPLICANT: Ishida, Yukako  
APPLICANT: Matsushima, Kouji  
APPLICANT: Kuno, Kouji  
TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same,  
TITLE OF INVENTION: pharmaceutical composition and method of immunologically analyzing human ADAMTS  
FILE REFERENCE: 057092  
CURRENT APPLICATION NUMBER: US/10/097,597  
PRIOR FILING DATE: 2002-03-15  
CURRENT APPLICATION NUMBER: 09/445,023  
PRIOR FILING DATE: 1999-12-03  
PRIOR APPLICATION NUMBER: JP 9-160422  
PRIOR FILING DATE: 1997-06-03  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 2  
LENGTH: 2184  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: exon  
LOCATION: (1)..(2184)  
US-10-097-597-2

Query Match 16.7%; Score 253; DB 9; Length 2184;

Best Local Similarity 60.9%; Pred. No. 8.2e-62;

Matches 412; Conservative 0; Mismatches 265; Indels 0; Gaps 0;

OY 783 GCGGCGGCGCGCGCTCCATCTCCCGCGCCAGGTGAGCTGCTTGGTGGCTG 842  
Db 20 GCGTTAGGAAGAGGATTTGTCTCAGCCGCCCTTATGTGGAACCAATGCTTGGCGAG 79  
OY 843 ACGCTCCATGCGCGGCTGTGATGCGCGGCGCTGACATTAATCTGCGACCTGGCCT 902  
Db 80 ACCAGTCATGCGCAAAATTCACGCGCATGTGCTTAAGCATTAATCTTCTACGTTGTTT 139  
OY 903 CCATGCCAATAGGCTGTACAGCCATGTAGCATGAGAACCAATCCGCTGGCGGTGG 962  
Db 140 CGGTGGCAGCGATTTGATCAAAACCCAGCATTCGTATTAATCACTTACCTTGGTGG 199  
OY 963 TGAAGTGTGTGTCTAGCGCAAGAGAGAGCGCTGGAATGAGCAAGACGCTGGCA 1022  
Db 200 TGAATCTTGGTCTATCCAGCATGAGAGAGGCGCGCAATGCACTCCCAATGCTGGCC 259  
OY 1023 CCACACTCAAGACTTTTTCAGAGTGGAGACACACACACAGCAGTGGAGATGACATG 1082  
Db 260 TCAGTCTGGGAACTTTTTCAGAGTGGAGAGACACACACACAGCAGTGGAGATG 319  
OY 1083 AGGACACTAGCATGACGATCTCTGTTACTGGGAGGATTTATGTGGCATTCATTCAT 1142

Db 320 CAGAGCACTATGACACACACATTCCTTTCACACAGACAGACTGTGTGGTCCAGACAT 379  
OY 1143 GTGACACCCCTGGGAATGGCAACAGTGGGACCATATGTCTTCAGAGCGCAGCTGTGTG 1202  
Db 380 GTGATACCTTGGGAGTGGCTGATGTGTGAACCTGTGTGATCCAGACAGAGAGCTGCTCG 439  
OY 1203 TGATTAAGAGAGAGCGCTCCACGAGCTTCACTGTGGCTCAGAAATCGACATTTAC 1262  
Db 440 TCATAGAAGATGATGTGTTTACAGATGTGCTTACACACAGCCCATGAATTAAGCCACG 499  
OY 1263 TTGGCTCTCCATGACGATTCCAAATTTCTGTGAAGAGACCTTTGGTTCACAGAAATA 1322  
Db 500 TTAACATGCCACATGATGATGCAAAAGCAGTGTGCCAGCCTTAATGTGTGAACAGGAT 559  
OY 1323 AGCGTTAATGTCTTCATCTCTTACACACATGATGATCTTAAGCCCTGGTCCAAATGCA 1382  
Db 560 CCCACATATGCGCTCAATGCTTTCACACGCGACACAGCAGCCTTGTCTCTTGA 619  
OY 1383 CTTCAGCCACCATCAGACAAATTCCTGATGATGGCCATGATGTTGGTGGACCTAC 1442  
Db 620 GTGCTTACATGATTAATCATTCATTTCTGATTAATGTCATGGGGAATGTTGATGGACA 679  
OY 1443 CACGAAGCAGATCCTG 1459  
Db 680 CTCAGAAATCCCATACAG 696

## RESULT 11

US-10-097-580-2

Sequence 2, Application US/10097580

Publication No. US20030032168A1

GENERAL INFORMATION:  
APPLICANT: Hirose, Kunitaka  
APPLICANT: Inoguchi, Eiji  
APPLICANT: Hakozaeki, Michinori  
APPLICANT: Ishioke, Keiko  
APPLICANT: Ishida, Yukako  
APPLICANT: Matsushima, Kouji  
APPLICANT: Kuno, Kouji  
TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical  
TITLE OF INVENTION: composition and method of immunologically analyzing human ADA  
FILE REFERENCE: 057092  
CURRENT APPLICATION NUMBER: US/10/097,580  
PRIOR FILING DATE: 2002-03-15  
CURRENT APPLICATION NUMBER: 09/445,023  
PRIOR FILING DATE: 1999-12-03  
PRIOR APPLICATION NUMBER: JP 9-160422  
PRIOR FILING DATE: 1997-06-03  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 2  
LENGTH: 2184  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: exon  
LOCATION: (1)..(2184)  
US-10-097-580-2

Query Match 16.7%; Score 253; DB 9; Length 2184;

Best Local Similarity 60.9%; Pred. No. 8.2e-62;

Matches 412; Conservative 0; Mismatches 265; Indels 0; Gaps 0;

OY 783 GCGGCGGCGCGCGCTCCATCTCCCGCGCCAGGTGAGCTGCTTGGTGGCTG 842  
Db 20 GCGTTAGGAAGAGGATTTGTCTCAGCCGCCCTTATGTGGAACCAATGCTTGGCGAG 79  
OY 843 ACGCTCCATGCGCGGCTGTGATGCGCGGCGCTGACATTAATCTGCGACCTGGCCT 902  
Db 80 ACCAGTCATGCGCAAAATTCACGCGCATGTGCTTAAGCATTAATCTTCTACGTTGTTT 139  
OY 903 CCATGCCAATAGGCTGTACAGCCATGTAGCATGAGAACCAATCCGCTGGCGGTGG 962

Db 140 CGGTGGCAGCCGATTTGACAAACACCCGACATTCGTAATTCAGTTAGCCGTGGCTGG 199  
Oy 963 TGAAGGTGGTGGTGGTGGGCGACAGACAGGCGTGAAGTGGAGCAAGAGCGTGGCA 1022  
Db 200 TGAAGATCTTGGTGCATCCACGATGACAGAGAGAGGCGGAGATGACCTCCAAATGCTGGCC 259  
Oy 1023 CCACACTGACAGACTTTTGCAGAGTGGCAGACACCAACAGAGCTGGAGATGACATG 1082  
Db 260 TCACCTCTGGGAACTTTTGCAGACTGGCAGAGACACCAACCCAGCTGACCGGATG 319  
Oy 1083 AGGACACTAGACATGACATCTCTGTTACTCGGAGGATTTAGTGGCATCATTCAT 1142  
Db 320 CAGACACTATGACACAGCAATCTTTTCACGACAGAGACTGTGTGGTCCAGACAT 379  
Oy 1143 GTGACACCCCTGGGATGGCAGAGCTTGACATATGTTCTCCAGAGCGCATCTGCTG 1202  
Db 380 GTGATACCTCTTGGATGGCTGATGTTGGACACTGTGTGATCCGAGCAGAACCTCTCG 439  
Oy 1203 TGATTTGAAGACAGATGGCTCCACGAGCGCTTGACTGTGGCTCAGCAAAATCGGACATTTAC 1262  
Db 440 TCATAGAAAGATGATGCTTTACAAAGCTGCCCTTCACCAAGCCCATGAAATTAAGCCAGTGT 499  
Oy 1263 TTGGGCTCTCCCATGACGATTTCCAAATTTCTGTGAAGACACCTTTGGTCCAGAGATA 1322  
Db 500 TTAAACATGCCACATGATGATGCAAGAGCAAGTGTGCCAGCTTAATGTTGTGAACCCAGAT 559  
Oy 1323 AGCGCTTAATGCTTCCATCTTACAGCATTTGATGATCTTAAGCCCTGGTCCAAATGCA 1382  
Db 560 CCCACATGATGGCGTCAATGCTTTCCAACTGGACACAGCCAGCTTGCTCTCTGCA 619  
Oy 1383 CTTCAGCCACCATCACAGAAATTCCTGATGATGGCCTGATGCTGTTGCTGACCTAC 1442  
Db 620 GTCCCTACATGATATCATCATTTCTGATTAATGCTCATGGGAAATGTTGATGACAAAGC 679  
Oy 1443 CACGAAGCAGATCCTG 1459  
Db 680 CTCAGAAATCCCATACAG 696

## RESULT 12

US-09-445-023A-2  
Sequence 2, Application US/09445023A  
Patent No. US20020119167A1  
GENERAL INFORMATION:  
APPLICANT: Hirose, Kunitaka  
APPLICANT: Inouuchi, Eiji  
APPLICANT: Hakezaki, Michinori  
APPLICANT: Ishioh, Keiko  
APPLICANT: Ishida, Yukako  
APPLICANT: Matsushima, Kouji  
APPLICANT: Kuno, Kouji  
TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical  
FILE REFERENCE: Q57092  
CURRENT APPLICATION NUMBER: US/09/445-023A  
CURRENT FILING DATE: 1999-12-03  
PRIOR APPLICATION NUMBER: JP 9-160422  
PRIOR FILING DATE: 1997-06-03  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 2  
LENGTH: 2184  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: exon  
LOCATION: (1)..(2184)  
US-09-445-023A-2

Query Match 16.7% Score 253; DB 10; Length 2184;  
Best Local Similarity 60.9%; Pred No. 8.2e-62;  
Matches 412; Conservative 0; Mismatches 265; Indels 0; Gaps 0;

Oy 783 GCGCGGCGCGCGCGCTCCATCTCCGCGCCCGCAGGTGGAGCTGCTGTGCTG 842  
Db 20 GCGTTAGGAAGAGGATTTGTGTCCAGCCCCCGTTATGTGGAACCAATGCTTGGCGAG 79  
Oy 843 AGCGCTCATGGCGGGTGTATGCGCCCGGCGCTCAGCATTTAGCTGTGACCTGGCT 902  
Db 80 ACCACTTCATGGCAGAAATTCACAGCGAGTGTCTTAACCATTTACCTTCTCAGTGTGTTT 139  
Oy 903 CCATGGCCAAATAGGCTGTACAGCCATGTAGCATGAGAACCAATCCGCTGGCGCTGG 962  
Db 140 CGGTGGCAGCCGATTTGACAAACACCCAGCATTCGTAATTCAGTTAGCTGGTGGTGG 199  
Oy 963 TGAAGTGGTGGTGTAGCGCAAGACAGACAGCTTGAAGTGAAGCAAGAGCTGCCA 1022  
Db 200 TGAAGATCTTGGTGCATCCACGATGAAAGAGGCGCGGAAGTGCATTCATGCTGCC 259  
Oy 1023 CCACACTGACAGACTTTTGCAGAGTGGCAGACACCAACACAGCTGGGAGATGACATG 1082  
Db 260 TCACCTCTGGGAACTTTTGCAGACTGGCAGAGACACCAACCCAGCTGACCGGATG 319  
Oy 1083 AGGACACTAGACATGACATCTCTGTTACTCGGAGGATTTATGAGCATCATTCAT 1142  
Db 320 CAGACACTATGACACAGCAATTTCTTTCCAGACAGAGACTTGTGTGGTCCAGACAT 379  
Oy 1143 GTGACACCCCTGGGATGGCAGAGCTTGACATATGTTCTCCAGAGCGCATCTGCTG 1202  
Db 380 GTGATACCTCTTGGATGGCTGATGTTGGAACTGTGTGATCCGAGCAGAACCTCTCG 439  
Oy 1203 TGATTTGAAGACAGATGGCTCCACGAGCGCTTGACTGTGGCTCAGCAAAATCGGACATTTAC 1262  
Db 440 TCATAGAAAGATGATGCTTTACAAAGCTGCCCTTCACCAAGCCCATGAAATTAAGCCAGTGT 499  
Oy 1263 TTGGGCTCTCCCATGACGATTTCCAAATTTCTGTGAAGACACCTTTGGTCCAGAGATA 1322  
Db 500 TTAAACATGCCACATGATGATGCAAGAGCAAGTGTGCCAGCTTAATGTTGTGAACCCAGAT 559  
Oy 1323 AGCGCTTAATGCTTCCATCTTACAGCATTTGATGATCTTAAGCCCTGGTCCAAATGCA 1382  
Db 560 CCCACATGATGGCGTCAATGCTTTCCAACTGGACACAGCCAGCTTGCTCTCTGCA 619  
Oy 1383 CTTCAGCCACCATCACAGAAATTCCTGATGATGGCCTGATGCTGTTGCTGACCTAC 1442  
Db 620 GTCCCTACATGATATCATCATTTCTGATTAATGCTCATGGGAAATGTTGATGACAAAGC 679  
Oy 1443 CACGAAGCAGATCCTG 1459  
Db 680 CTCAGAAATCCCATACAG 696

## RESULT 13

US-09-918-171A-8  
Sequence 8, Application US/09918171A  
Patent No. US20020110894A1  
GENERAL INFORMATION:  
APPLICANT: Apte, Suneel  
APPLICANT: Hurskainen, Taina L.  
APPLICANT: Hirohata, Satoshi  
TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases  
FILE REFERENCE: 26473/04193  
CURRENT APPLICATION NUMBER: US/09/918-171A  
CURRENT FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 09/369,364  
PRIOR FILING DATE: 1999-08-06  
NUMBER OF SEQ ID NOS: 31  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 8  
LENGTH: 3638  
TYPE: DNA  
ORGANISM: Mus musculus ADAMTS-8  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (278)..(2992)

NAME/KEY: misc\_feature  
LOCATION: (3636) \*  
OTHER INFORMATION: n = T  
US-09-918-171A-8

Query Match 15.6%; Score 237; DB 10; Length 3638;  
Best Local Similarity 61.9%; Pred. No. 3,2e-57;  
Matches 394; Conservative 0; Mismatches 240; Indels 3; Gaps 1;

QY 804 TCTCCGCGGCCCCGAGCTGAGCTGCTTCTGCTGCTGAGCCGCTCCATGCGCGGTGT 863  
DB 966 TGTCCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1025  
QY 864 ATGCGCGGCGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 923  
DB 1026 ATGCGCGGCGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1085  
QY 924 GCGATGCTAGCTGAGCAACCATCCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 983  
DB 1086 AGCACCAGCATCAGAACTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1145  
QY 984 ACAAGCAGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1043  
DB 1146 AAGAAAGTGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1205  
QY 1044 AGTGGCAGCAGCAACCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1103  
DB 1206 GCTGGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1265  
QY 1104 TCTGCTTACTGCGAGAGATTTATGTGGGCA--TCATTCATGTGACACCTGGGAATG 1160  
DB 1266 TCTGCTTACTGCGAGAGATTTATGTGGGCA--TCATTCATGTGACACCTGGGAATG 1325  
QY 1161 CAGAGCTGGAGCATATGCTTCTCAGAGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1220  
DB 1326 CAGAGCTGGAGCATATGCTTCTCAGAGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1385  
QY 1221 TCCAGCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1280  
DB 1386 TGCAGGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1445  
QY 1281 ATTCCAAATTTGTGAGAGACCTTTGCTTCACAGAAATTAAGCGCTTAATGTCTTCCA 1340  
DB 1446 ATTCCAAATTTGTGAGAGACCTTTGCTTCACAGAAATTAAGCGCTTAATGTCTTCCA 1505  
QY 1341 TCCCTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1400  
DB 1506 TCCCTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1565  
QY 1401 AATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1437  
DB 1566 AGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1602

RESULT 14  
US-09-918-171A-3  
Sequence 3, Application US/09918171A  
Patent No. US20020110894A1  
GENERAL INFORMATION:  
APPLICANT: Apte, Suneel  
APPLICANT: Hurskainen, Tiina L.  
APPLICANT: Hirohata, Satoshi  
TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases  
FILE REFERENCE: 26473/04193  
CURRENT APPLICATION NUMBER: US/09/918,171A  
CURRENT FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 09/369,364  
PRIOR FILING DATE: 1999-08-06  
SOFTWARE: PatentIn Ver. 2.1  
SD ID NO 3  
LENGTH: 1520  
TYPE: DNA

ORGANISM: homo sapiens ADAMTS-5  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (296)  
OTHER INFORMATION: n=T  
NAME/KEY: misc\_feature  
LOCATION: (1443)  
OTHER INFORMATION: n=A  
NAME/KEY: misc\_feature  
LOCATION: (1457)  
OTHER INFORMATION: n=G  
NAME/KEY: misc\_feature  
LOCATION: (1458)  
OTHER INFORMATION: n=T  
US-09-918-171A-3

Query Match 13.9%; Score 210.2; DB 10; Length 1520;  
Best Local Similarity 98.6%; Pred. No. 9.9e-50;  
Matches 212; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1253 GCACATTTACTTGCGCTCTCCCATGAGCATTCGCAATTCGTGAAAGACCTTGCTTC 1312  
DB 1 GCACATTTACTTGCGCTCTCCCATGAGCATTCGCAATTCGTGAAAGACCTTGCTTC 60  
QY 1313 ACAGAGATAGCGCTTATGCTTCCATCCCTACACATTTGATGATGATGATGATGATG 1372  
DB 61 ACAGAGATAGCGCTTATGCTTCCATCCCTACACATTTGATGATGATGATGATGATGATG 120  
QY 1373 TCCAATGCACTTACGCGCAGCATCAGCAATTCCTGATGATGATGATGATGATGATG 1432  
DB 121 TCCAATGCACTTACGCGCAGCATCAGCAATTCCTGATGATGATGATGATGATGATGATG 180  
QY 1433 CTGAGCTTACGCGCAGCAATTCCTGCGCGGGA 1467  
DB 181 CTGAGCTTACGCGCAGCAATTCCTGCGCGGGA 215

RESULT 15  
US-09-918-171A-10  
Sequence 10, Application US/09918171A  
Patent No. US20020110894A1  
GENERAL INFORMATION:  
APPLICANT: Apte, Suneel  
APPLICANT: Hurskainen, Tiina L.  
APPLICANT: Hirohata, Satoshi  
TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases  
FILE REFERENCE: 26473/04193  
CURRENT APPLICATION NUMBER: US/09/918,171A  
CURRENT FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 09/369,364  
PRIOR FILING DATE: 1999-08-06  
NUMBER OF SEQ ID NOS: 31  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 10  
LENGTH: 739  
TYPE: DNA  
ORGANISM: Homo sapiens ADAMTS-8  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (3)..  
US-09-918-171A-10

Query Match 13.5%; Score 205; DB 10; Length 739;  
Best Local Similarity 60.4%; Pred. No. 2.4e-48;  
Matches 356; Conservative 0; Mismatches 230; Indels 3; Gaps 1;

QY 797 CGCTTCATCTCCGCGCGCCAGCTGAGCTGCTTGTGTGCTGAGCGCTCCATGCGCG 856  
DB 60 CGGTTGTGTCTGAGCGCGCGCTGCTGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 119  
QY 857 CGGTTGTGTGCGCGCGCGCTGAGCATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 916  
DB 120 GCTTCTACGCGCGCGAGCTGAGCAATTCCTGAGCATTCCTGAGCATTCCTGAGCATTCCT 179

```
OY 917 CTGTACAGCCATGCTAGCATGAGAACCATCCGCTGGCGTGTGAGGTGGTG 976
    ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 180 ATCTACAAAGCACCACGATGAGAAATTCATCAACCTGATGTGTAAAGTGTGATC 239
OY 977 CTAGCCGACAAAGGACAAAGACCTGGAAGTGAACAAGAACCTGCCACCACTCAAGAAC 1036
    ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 240 GTAGAAATGAAAAATGGGGCCAGAGGTGTCCCAAAATGGGGGCTTACACTGCTAAC 299
OY 1037 TTTTCAGATGGCAGCAGCAACCAACAGCTGGGAGATGACCATGAGAGCACTAGCAT 1096
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 300 TTCTGCAACTGGCAGCGGGGCTTTCACCAAGCCAGCAGCCGACCCAGAGCACTACGAC 359
OY 1097 GCAGCTATCTGTCTTACTGCGGAGGATTATGTGGGATCA--TTCAATGACACCCCTG 1153
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 360 ACGGCATCTGCTCAGCAGACAGAACTTGTGTGGGAGAGGGGCTGTGTACACCCCTG 419
OY 1154 GGAATGGCAGAGCTTGGGACCATATGTCTCCAGAGCCGAGCTGTGTGATTGAAGAC 1213
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 420 GGTGTGGCAGACATGGGACCATTTGTGACCCCAAAAGCTCCTCCGTATCGAGGAT 479
OY 1214 GATGGCTTCAGAGCGAGCTTTCACCTGTGGCTACGAAATCGGACATTTACTGGCCCTCC 1273
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 480 GAGGGGCTCCAGCGCGGCCACACCTGGCCCATGAACTAGGGCAGCTCCTCAGCATGCC 539
OY 1274 CATGACGATTCGAAATTCGTGTGAAGAGACCTTGGTCCACAGAAATAGCGCTTAATG 1333
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 540 CACGAGCACTCCAGCGCCTGACACAGGGCTTGGGGCCCATGGGCAAGCAGCAGTGATG 599
OY 1334 TCTTGCATCCTTACCAAGCATTCATCATCTAAGCCCTGTGCAAAATGCA 1382
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 600 GCACGGCTGTGTCTCACCCTGAACCAAGAGCGTGGCTGTCCCTCTGCA 648
```

Search completed: April 26, 2003, 09:50:41  
Job time : 365.637 secs

GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 25, 2003, 16:50:57 ; Search time 4152.63 Seconds  
(without alignments)  
5300.174 Million cell updates/sec

Title: US-10-050-200-1

Perfect score: 1359  
Sequence: 1 gaatcgccatgctccagac.....atgacaaagggtagctgcac 1359

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estlba:\*  
2: em\_estlhum:\*  
3: em\_estlin:\*  
4: em\_estlmu:\*  
5: em\_estrov:\*  
6: em\_estrpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_estl1:\*  
10: gb\_estl2:\*  
11: gb\_hlc:\*  
12: gb\_estl3:\*  
13: gb\_estl4:\*  
14: gb\_estl5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	577.4	42.5	598	13	BM069515
c 2	400.6	29.5	681	12	BE912007
c 3	374	27.5	440	12	BF850604
c 4	346.2	25.5	420	10	AM353152
c 5	332.2	24.4	700	13	BI691035
c 6	313.4	23.1	835	13	BI602707

Result No.	Score	Query Match	Length	DB ID	Description
c 7	299.2	22.0	344	12	BF850605
c 8	276	20.3	832	11	BC540485
c 9	273.2	20.1	2931	12	BC003269
c 10	264.4	19.5	327	12	BF850549
c 11	231.8	17.1	573	14	BQ353240
c 12	223	16.4	645	9	AL712592
c 13	214	15.7	274	10	AM357906
c 14	210.8	15.5	452	12	BF853210
c 15	192	14.1	574	13	BM069747
c 16	181	13.3	213	10	AM353150
c 17	179	13.2	680	12	BF463232
c 18	178.8	13.2	652	9	AI400905
c 19	178	13.1	602	10	AM137467
c 20	176	13.0	600	9	AL603427
c 21	163.8	12.1	566	10	AV617111
c 22	158.2	11.6	667	10	BB624639
c 23	152.8	11.2	613	10	BB612585
c 24	148.6	10.9	660	10	BB651049
c 25	148.6	10.9	662	10	BB642806
c 26	147	10.8	464	9	AI969885
c 27	147	10.8	528	9	AA288689
c 28	146	10.7	544	12	BF078689
c 29	144.6	10.6	565	10	BB610011
c 30	139.8	10.3	431	9	AI596966
c 31	139.8	10.3	495	9	AI674956
c 32	139.8	10.3	580	9	AA530477
c 33	139	10.2	468	10	BB646878
c 34	136.2	10.0	543	9	AA635657
c 35	136	10.0	654	10	BE016461
c 36	134.4	9.9	377	9	AI464076
c 37	132.8	9.8	661	10	BB220373
c 38	131.2	9.7	777	10	BB641727
c 39	129.4	9.5	678	10	BB655188
c 40	120	8.8	1025	17	CNS029MR
c 41	118.2	8.7	636	12	BF707715
c 42	116.6	8.6	630	10	BB611853
c 43	116.6	8.6	631	10	BB658345
c 44	112.8	8.3	354	9	AA855532
c 45	110.6	8.1	752	10	BE273010

## ALIGNMENTS

RESULT 1  
BM069515/c  
LOCUS  
DEFINITION  
BM069515  
CDNA clone IMAGE:5674169 3' similar to TR:075173 KIA00688  
PROTEIN. (1) ; mRNA sequence.  
ACCESSION  
BM069515  
VERSION  
BM069515.1 GI:16909594  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
1 (bases 1 to 598)  
Melton, D., Brown, J., Kently, G., Penmutt, A., Lee, C., Kaestner, K., Lemishka, I., Searce, M., Brestell, J., Gradwohl, G., Clifton, S., Hillier, L., Maria, M., Pape, D., Wylie, T., Martin, J., Blistain, A., Schmitt, A., Theising, B., Riller, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagarelisvili, R., Williams, T., Jackson, Y., and Bowers, Y.  
Endocrine Pancreas Consortium  
Unpublished (2000)  
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557





[illegible]



REFERENCE 1 (bases 1 to 835)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Miklos Palokovits, M.D., Ph.D.  
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki Toshiyuki and Piero Carninci (RIKEN)  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMNL at:  
<http://image.lnl.gov>  
 Plate: LMNL1769 row: e column: 13  
 High quality sequence stop: 831.  
 Location/Qualifiers  
 1. 835  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5303796"  
 /clone\_lib="NIH-MGC\_96"  
 /tissue\_type="hypothalamus"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcagag ) ; Oligo-dt primed using primer 5'-TTTTTTTTTTTTTTTNN-3', size-selected for average insert size 2.3 kb and normalized for ROP 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH-MGC Library."  
 BASE COUNT 186 a 243 c 267 g 139 t  
 ORIGIN  
 Query Match 23.1%; Score 313.4; DB 13; Length 835;  
 Best Local Similarity 93.6%; Pred. No. 7.3e-62;  
 Matches 381; Conservative 0; Mismatches 21; Indels 5; Gaps 5;

QY 7 GCCATGTCCCAAGAGCGTCGATCCCGGAGAGCGCTTGGCGGCGCTGGC-TGTGGGG 65  
 |||||  
 DB 429 GCCATGTCCCAAGATAGCGTCGATCCCGGAGAGCGCTTGGCGGCGCTTGGGG 488  
 QY 66 AGCCCAACCGCTCTCGTCCCGCATTTGTCGCTGCTGGTGGTGGCTGCTTCT 125  
 |||||  
 DB 489 AGCCCAACCGCT 548  
 QY 126 GCTACTGTGGGCT 185  
 |||||  
 DB 549 GCTACTGTGGGCT 608  
 QY 186 GGAGATCGGTTTCCAGAGAACTCAAGCGGAGCGTCTCGCTGGCTGGGACCCCTTC 245  
 |||||  
 DB 609 GGAGATCGGTTTCCAGAGAACTCAAGCGGAGCGTCTCGCTGGCTGGGACCCCTTC 667  
 QY 246 CAGGCTGTGGTCCCTTGGAGAGAGCGCTGCTACTAG-AGCTGGAGCAGG 304  
 |||||  
 DB 668 CAGGCTGTGGTCCCTTGGAGAGAGCGCTTGGAGAGAGCGCTGCTACTAGTGAAGCTGGAGCAGG 727  
 QY 305 ACTCGGTTGTGAGGTCGAGGGGCTGACAGTGCAGTACCTGGGCGGCGCTGAGCTGC 364  
 |||||  
 DB 728 ACTCGGTTGTGAGGTCGAGGGGCTGACAGTGCAGTAACTGGGCGGCGCTGAGCTTG 787  
 QY 365 TGGGTGGAGACAGCGCTGACACTTACTGCTGAGCAGCATCATGGA 411  
 |||||  
 DB 788 T-GGTGGAGACAGCGCTG-ACCTACCTGACTGGAACAATCAATGGA 832

RESULT 7  
 BF850605/c 344 bp mRNA linear EST 16-JAN-2001  
 LOCUS BF850605

DEFINITION PMI-EN0065-191100-007-f12 EN0065 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BF850605  
 VERSION BF850605.1 GI:12237767  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM human.  
 REFERENCE 1 (bases 1 to 344)  
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.B., Carvalhal, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.  
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 MEDLINE 20202663  
 COMMENT Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: [asimpson@ludwig.org.br](mailto:asimpson@ludwig.org.br)  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
<http://www.ludwig.org.br/scripts/gethtml2.pl?l1=PMI&l2=PMI-EN0065-191100-007-f12&l3=2000-11-19&l4=1>  
 Seq.primer: puc 18 forward  
 High quality sequence start: 6  
 High quality sequence stop: 344.  
 Location/Qualifiers  
 1. 344  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="EN0065"  
 /dev\_stage="Adult"  
 /note="Organ: lung\_normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent Application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."  
 BASE COUNT 91 a 89 c 100 g 64 t  
 ORIGIN  
 Query Match 22.0%; Score 299.2; DB 12; Length 344;  
 Best Local Similarity 97.0%; Pred. No. 8.9e-59;  
 Matches 326; Conservative 0; Mismatches 8; Indels 2; Gaps 2;

QY 938 TTGACACAGCCATTCTGTTACCGCTCAGAGACCTGTGTGGAGTCTCCACTTGGACAGCC 997  
 |||||  
 DB 344 TTGACACAGCCATTCTGTTACCGCTCAGAGACCTGTGTGGAGTCTCCACTTGGACAGCC 285  
 QY 998 TGGGTATGGCTATATGTGGGACACCGTCTGTGACCCGCTCGAGAGCTGTGCCATTTGGAGG 1057  
 |||||  
 DB 284 TGGGTATGGCTATATTTGGGACACCGTCTGTGACCCGCTCGAGAGCTGTGCCATTTGGAGG 225  
 QY 1058 ATGATGGGCTGC-AGTCAAGCTTCACTGTGCTCATGAACTGGGTGATGTTCAAACATG 1116  
 |||||  
 DB 224 ATGATGGGCTTCTTAATCAAGCTTCACTGTGCTCATGAACTGGGTGATGTTCAAACATG 165  
 QY 1117 CTCATGACAACTCCAAAGCATGATC-AGTTTGAATGGGCTTTGAGCAGCCCTTCGCCA 1175  
 |||||  
 DB 164 CTCATGACAACTCCAAAGCATGATCAGTTTGAATGGGCTTTGAGCAGCCCTTCGCCA 105  
 QY 1176 TGTCAATGGCCCTGTGATGGCTCATGTGATCTTGAAGAGACCCCTGCTCCCTGAGATGC 1235  
 |||||  
 DB 104 TGTCAATGGCCACTGTGATGGCTCATGTGATCTTGAAGAGACCCCTGCTCCCTGAGATGC 45

QY	1236	CCGCTTCATCACTGACTGCTCTGGACAATGGCGATGG	1271
Db	44	CCGCTTCATCACTGACATGCTCTGGACAATGGCGATGG	9
RESULT	8		
LOCUS	BC540485		
DEFINITION	BC540485	832 bp	mRNA
	602566210F01 NIH_MGC_77	Homo sapiens	CDNA clone IMAGE:4693387 5',
			mRNA sequence.
ACCESSION	BC540485		
VERSION	BC540485.1	GI:13532718	
KEYWORDS	EST.		
SOURCE	human		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.		
REFERENCE	1 (bases 1 to 832)		
AUTHORS	NIH-MGC <a href="http://mgc.ncl.nih.gov/">http://mgc.ncl.nih.gov/</a> .		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D.		

Plate: BLCM1516 row: d column: 04  
High quality sequence stop: 686.  
\* Location/Qualifiers

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image:4693587"
/clone_id="NH_MGC_77"
/lab_host="DH10B (TI phage-resistant)"
/notes="Organ: lung; Vector: pDMR-LIB (Clontech); Site:1:
SfiI (ggcgccctggcgc); Site_2: SfiI (ggccatattggc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGCCCATTAATGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGCGCCGACACATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.9
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NH_MGC Library."
```

Query Match	20.3%	Score 276;	DB 12;	Length 832;
Best Local Similarity	85.1%;	Pred. No. 3.2e-53;		
Matches 344;	Conservative 55;	Mismatches 0;	Indels 5;	Gaps 3;

OY	7	GCCATGCTCCCAAGACAGGCTGCGATCTCCCGGAGGGGCTTGGCAGGGCGCTGTGGGGGA	66
Db	431	GCCATGTCCCAAGACAGGCTGCATCCCGGAGGAGGCTTGGCAGGGCGCTGGCTGTGGGGA	490
OY	67	GCCCAACCCCTGCTCTCTGCTCCCAATTGTGCGGCTCTCTCTGGTGTGTGGCTGCTTCTG	126
Db	491	GCCCAA - CCTGCGCTCTCTCTCCCAATTGTGCGGCTCTCTGGCTGTGTGGCTGCTTCTG	549
OY	127	CTACTGCTGGGCTCTCTCTCTGCTGCTCAAGCCGGGCTGGGCAAGCCCTTCCCGGGAGAG	186
Db	550	CTACTGCTGGGCTCTCTCTCTGCTGCTCAAGCCGGGCTGGGCAAGCCCTTCCCGGGAGAGG	609
OY	187	GAGATCGTGTTCAGAGAAAGCTCAAGGCAAGCGTGGCTGGGCTGTGGGGCAAGCCCTGCC	246
Db	610	GAGATCG - GTTTCAGAGAAAGCTCAAGGCGCGCTGCTGGCTGTGGCTGGGGGCGCTCTCC	668

[illegible]

RESULT	9
BC003269	
LOCUS	BC003269
DEFINITION	Mus musculus. Similar to a disintegrin-like and metalloprotease (repsilysin type) with thrombospondin type 1 motif, 1, clone IMAGE:3499206, mRNA.
ACCESSION	BC003269
VERSION	BC003269.1
KEYWORDS	GI:14707747
SOURCE	HTC
	house mouse.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus,  
 1 (bases 1 to 2931)  
 Strausberg, R.  
 Direct Submission  
 Submitted (20-FEB-2001) National Institutes of Health, Mammalian  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

REMARK	NIH-MGC Project URL: <a href="http://mgc.ncl.nih.gov">http://mgc.ncl.nih.gov</a>
COMMENT	Contact: MGC help desk

Tissue Procurement: Jeffrey Green M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: amgebcm.tmc.edu  
Gunnarline, P.H., Garcia, A.M., Lu, X., Huylk, S.W., Hale, S.M.,  
Yoon, V.S., Kowls, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,  
Richards, S., Gibbs, R.A.

Clone distribution: MCC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
 Series: IRAK Plate: 9 Row: b Column: 7  
 This clone has the following problem: frame shifted.

FEATURES	Location/Qualifiers
source	1. , 2931

```

/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="IMAGE:3499206"
/tissue-type="Mammary tumor. C3(1)-Tag model. Infiltrating
ductal carcinoma. 5 month old virgin mouse."

```

	/note--"Vector: PCMV-SPORT6"
BASE COUNT	708 a 812 c 811 g 600 t
ORIGIN	.

Query Match	20.1%	Score 273.2	DB 11	Length 2931
Best Local Similarity	61.7%	Pred. No. 2.7	5E-52	
Matches 453; Conservative	0;	Mismatches 278;	Indels 3;	Gaps 1;
Oy	620	GCCTCCAGACCCCGAAGAGCCAGCGCTTTCCTCACTAGTAGATTTTGTGGAGACACTGG	679	
db	1124	GACCCGGAAGATAGGAGGAAGAGGACATTTGTCTCCAGCCGCCGTTATTTGTGGAACCATATGC	1183	

```

QY 680 TGGTGGCAGATGACAGATGCGCCGATTCACAGTGGCGGGCTAAAGCGCTACTGCTAA 739
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1184 TGGTGGCTACACAGTGCATGGCGGACTTCACAGGAGCGCTCTAAAGCATCTTCTA 1243
QY 740 CAGTGTGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 799
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1244 CCTGTGTTGGTGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1303
QY 800 TGGTGGTGCATGGCTGAGTGCATGGCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 859
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1304 TGGTGGTGGTGAAGATCTTGTGCTATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1363
QY 860 GTGCTGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 919
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1364 ATGACAGCTCTACACCTTCGGAATTTCTGCACTGGCAGAAACACACACAGCCCAAGT 1423
QY 920 ACTGGAGACCTTACACCTTTCAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 979
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1424 ACCGGATCCAGACAGCAGTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1483
QY 980 TCTCCACTTGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1039
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1484 CCCACACGCTGACACTCTCGGATGGCAGATGTTGAAGTGTAGACCCAGCAGAGA 1543
QY 1040 GCTGTGCCATTTGAGAGATGATGGCTCCAGTCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1099
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1544 GCTGCTCACTATAGAAAGATGATGTTTGCAGAGCCGCTTACACAGCAGCAGCAGATG 1603
QY 1100 GTCATGTCTTCAACATGCTTCCATGACACTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1159
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1604 GGCATCTGTTTAACTACGCGCAGAGATGATGACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1660
QY 1160 TGAGACCTCTCTGCGCATGATGAGCCCTGTGATGGCTCATGATGATGATGATGATGATGATGAT 1219
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1661 TGACTGGCATCTCTATCTGATGAGCCTGATGCTTCCAGCTTACAGCAGCAGCAGCAGCAGCAG 1720
QY 1220 GGTCCCTGAGAGTGGCGCCCTTCATCATGACTGACTTCTTGAGACATGGCTATGGGCACTGTC 1279
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1721 GGTACCTGTCAGTGGCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1780
QY 1280 TCTTACACAAACAGAGGCTTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1339
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1781 TATATGACAAAGCCCGAGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1840
QY 1340 ATGACAAAGGGTAG 1353
    ||| ||| ||| |||
Db 1841 ATGCCAACGCCAG 1854
    ||| ||| ||| |||

```

```

RESULT 10
BF850549/c 337 bp mRNA linear EST 16-JAN-2001
DEFINITION PM1-EN0065-191100-006-a01 EN0065 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF850549
VERSION BF850549.1 GI:12237711
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 327)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shogun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.

```

---

```

FEATURES
    source
        1..327
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_lib="EN0065"
            /dev_stage="Adult"
            /note="Organ: Lung normal; Vector: puc18; Site:1: SmaI;
            Site:2: SmaI; A mini-library was made by cloning products
            derived from ORESTES PCR (U.S. Letters Patent application
            No. 196,716 - Ludwig Institute for Cancer Research)
            profiles into the puc 18 vector. Reverse transcription of
            tissue mRNA and cDNA amplification were performed under
            low stringency conditions."
            BASE COUNT      84 a      78 c      97 g      68 t

```

---

```

Query Match      19.5%; Score 264.4; DB 12; Length 327;
Best Local Similarity 97.6%; Pred. No. 9,6e-51;
Matches 279; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

```

---

```

QY 1035 TGGAGCTGTGCGATGTGAGAGATGATGGCTCCAGTACGCTTACGCTGCATCA 1094
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 327 TCGGACCTGTCCATGTTGAGAGATGATGGCGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 268
QY 1095 ACTGGTCACTGCTTCACATGCTCCATGACAGCAACTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1154
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 267 ACTGGTATATGCTTCAACATGCTCCATGACAGCAACTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 208
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1155 GCGTTGAGCAGCCTCGCCATGTCATGAGCGCCCTGTGATGGCTCATGATGATGATGATGATGATGAT 1214
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 207 GCGTTGAGCAGCCTCGCCATGTCATGAGCGCCCTGTGATGGCTCATGATGATGATGATGATGATGAT 148
QY 1215 GCGCTGTGCTCCCTGAGTGGCGCGCTTCACTACATGCTTCTTGAGACATGGCTATGAGCA 1274
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 147 GCGCTGTGCTCCCTGAGTGGCGCGCTTCACTACATGCTTCTTGAGACATGGCTATGAGCA 88
QY 1275 CTTGTCTTACAAACAGAGGCTTCATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 87 CTGTCTTACAAACAGAGGCTTCATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 43

```

---

```

RESULT 11
B0353240/c 573 bp mRNA linear EST 20-MAY-2002
DEFINITION PM2-HT0927-290900-005-h09 HT0927 Homo sapiens cDNA, mRNA sequence.
ACCESSION B0353240
VERSION B0353240.1 GI:21017296
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 573)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

```

**TITLE** Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

**JOURNAL** Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

**MEDLINE** 20202663

**COMMENT** Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?c1=PM2&c2=PM2-HT0927-290900-005-h09&c3=2000-09-29&c4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 571.

**FEATURES**  
source  
1..573  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone.lib="HT0927"  
/dev\_stage="Adult"  
/note="Organ: head,neck; Vector: puc18; Site\_1: Sma1; Site\_2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

**BASE COUNT** 109 a 170 c 182 g 112 t

**ORIGIN**

Query Match 17.1%; Score 231.8; DB 14; Length 573;  
Best Local Similarity 98.5%; Pred. No. 4.3e-43;  
Matches 255; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 7 GCCATGTCACAGACAGCGTCGATCCCGGAGGCGCTTGGCAGGCGCTGTGGGGA 66  
DB 270 GCCATGTCACAGACAGCGTCGATCCCGGAGGCGCTTGGCAGGCGCTGTGGGGA 211  
QY 67 GCCAACCTGCGCTCTCTCTCCCATTTGTCGCGCTCTCTGCTGTGGTGTCTTG 126  
DB 210 GCCAACCTGCGCTCTCTCTCCCATTTGTCGCGCTCTCTGCTGTGGTGTCTTG 151  
QY 127 CTACTGCTGCGCTCTCTCTCTGCGCTCAGCCCGGCTGAGCGCCGCTCCCGGGAGGA 185  
DB 150 CTACTGCTGCGCTCTCTCTCTGCGCTCAGCCCGGCTGAGCGCCGCTCCCGGGAGGA 91  
QY 186 GGAGATGCTGTTTCCAGAGAGCTCAACGGCAGCGTCTGCTGGCTCGGGACCCCTG- 244  
DB 90 GGAGATGCTGTTTCCAGAGAGCTCAACGGCAGCGTCTGCTGGCTCGGGACCCCTG 31  
QY 245 CCAGGCTGTTTGGCCGCTT 263  
DB 30 CCAGGCTGTTTGGCCGCTT 12

**RESULT 12** 645 bp mRNA linear EST 22-MAR-2002  
LOCUS AL712592  
DEFINITION DKFZP686A228.F1 686 (synonym: hlc03) Homo sapiens cDNA clone  
ACCESSION AL712592  
VERSION AL712592.1 GI:19695948  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE 1 (bases 1 to 645)  
Poustka,A., Wellenreuther,R., Mewes,H.W., Well,B. and Wiemann,S.).

**TITLE** EST (Poustka,A., Wellenreuther,R., Mewes,H.W., Well,B. and Wiemann,S.)

**JOURNAL** Unpublished (1999)

**MEDLINE** Contact: Poustka A.J.

**COMMENT** Department Lehnach  
Max-Planck-Institute for Molecular Genetics  
Innistrasse 73, 14195 Berlin, Germany  
Tel: +49-30-84131623  
Fax: +49-30-84131128  
Email: poustka@mpimg-berlin-dohlem.mpg.de  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.  
No sl. sequence available.  
This clone (DKFZP686A228) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

**FEATURES**  
source  
1..645  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone.lib="DKFZP686A228"  
/clone.lib="686 (synonym: hlc03)"  
/tissue\_type="human skeletal muscle"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/note="Vector: pTriplex2; Site\_1: SfiIA; Site\_2: SfiIB; cDNA collection"

**BASE COUNT** 153 a 194 c 198 g 100 t

**ORIGIN**

Query Match 16.4%; Score 223; DB 9; Length 645;  
Best Local Similarity 100.0%; Pred. No. 5e-41;  
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GCCATGTCACAGACAGCGTCGATCCCGGAGGCGCTTGGCAGGCGCTGTGGGGA 66  
DB 423 GCCATGTCACAGACAGCGTCGATCCCGGAGGCGCTTGGCAGGCGCTGTGGGGA 482  
QY 67 GCCAACCTGCGCTCTCTCTCCCATTTGTCGCGCTCTCTGCTGTGGTGTCTTG 126  
DB 483 GCCAACCTGCGCTCTCTCTCCCATTTGTCGCGCTCTCTGCTGTGGTGTCTTG 542  
QY 127 CTACTGCTGCGCTCTCTCTCTGCGCTCAGCCCGGCTGAGCGCCGCTCCCGGGAGGAG 186  
DB 543 CTACTGCTGCGCTCTCTCTCTGCGCTCAGCCCGGCTGAGCGCCGCTCCCGGGAGGAG 602  
QY 187 GGAGATGCTGTTTCCAGAGAGCTCAACGGCAGCGTCTGCTGGCTCGGGACCCCTG 229  
DB 603 GGAGATGCTGTTTCCAGAGAGCTCAACGGCAGCGTCTGCTGGCTCGGGACCCCTG 645

**RESULT 13** 274 bp mRNA linear EST 25-APR-2001  
LOCUS AW357906  
DEFINITION 41683 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.  
ACCESSION AW357906  
VERSION AW357906.1 GI:6861912  
KEYWORDS EST.  
SOURCE cow.  
ORGANISM Bos taurus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.  
REFERENCE 1 (bases 1 to 274)  
Smith,T.P.L., Grose,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.M., Rohrer,G.A., Chitko-Mckown,C.G., Pettea,G., Holt,L., Karayancheva,S., Liang,F., Quackenbush,J. and Keefe,J.W.



TITLE  
JOURNAL  
COMMENT

Hillier, L., Marie, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,  
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas,  
M., Gibbons, M., McCann, R., Cole, R., Tsagarelis, R., Williams, T.,  
Jackson, Y., and Bowers, Y.  
Endocrine Pancreas Consortium  
Unpublished (2000)  
Other ESTs: 1e88509.x1  
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmelton@bioh.harvard.edu  
Library was constructed by Dr. Douglas Melton DNA sequencing by:  
Washington University Genome Sequencing Center For information on  
obtaining a clone please contact: Juliana Brown  
(brownjfas.harvard.edu) This sequence now available from the IMAGE  
consortium, for clone orders contact: ImageImage.llnl.gov  
High quality sequence stop: 440.

FEATURES  
source

1..574  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone IMAGE:5674169"  
/clone\_lib="Melton Normalized Human Islet 4 N4-HIS 1"  
/sex="Both"  
/tissue\_type="Islets of Langerhans"  
/dev\_stage="Adult"  
/lab\_host="DH10B"  
/note="Organ: Pancreas; Vector: pSPORT1; Site\_1: Not 1;  
Site\_2: Sal 1; Starting library constructed using  
Superscript Plasmid Library kit (Life Technologies). cDNA  
made by oligo-dT priming. Size-selected by column  
fractionation: average insert size 1.08 kb. Library was  
amplified once on solid support and plasmid DNA from  
library was prepared. The library DNA was normalized by  
method #4 from Bonaldo, Lennon, and Soares 1996 Genome  
Research 6:791-806; 0.5 microgram single-stranded library  
plasmid DNA was mixed with 5 micrograms PCR product  
representing library inserts and hybridized to an EcoT of  
20. Single-stranded (unhybridized) plasmids were isolated  
by hydroxyapatite chromatography and used to make this  
library."

BASE COUNT 130 a 168 c 179 g 97 t  
ORIGIN

Query Match 14.1%; Score 192; DB 13; Length 574;  
Best Local Similarity 100.0%; Pred. No. 6.8e-34;

Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

0Y 7 GCCATGTCACAGAGGCTCCGAGAGGCTTGGACGAGCGCTGTGGGA 66  
DB 383 GCCATGTCACAGAGGCTCCGAGAGGCTTGGACGAGCGCTGTGGGA 442  
0Y 67 GCCAACCCTGCTCTGCTCCCATGTGCGGCTCTCTGCTGTGCTCTTG 126  
DB 443 GCCCAACCCTGCTCTGCTCCCATGTGCGGCTCTCTGCTGTGCTCTTG 502  
0Y 127 CTACGCTGCGCTCTCTGCTCCCATGTGCGGCTCTCTGCTGTGCTCTTG 186  
DB 503 CTACGCTGCGCTCTCTGCTCCCATGTGCGGCTCTCTGCTGTGCTCTTG 562  
0Y 187 GAGATCGTGT 198  
DB 563 GAGATCGTGT 574

Search completed: April 26, 2003, 04:35:46  
Job time : 4171.63 secs





```
|||||
Db 536 CTACTGCTGGCTCTCTCTGCCCCCTACAGCCCGGCTGGCCAGCCCTCCCGGGAGAG 595
Oy 187 GAGATGCTGTTTCCAGAGAAAGTCAAGGCGACGCTGCTGCTGCTGGGCAACCTGCG 246
Db 596 GAGATGCTGTTTCCAGAGAAAGTCAAGGCGACGCTGCTGCTGCTGGGCAACCTGCG 655
Oy 247 AGGCTGTTGTCGCGCTTGGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 306
Db 656 AGGCTGTTGTCGCGCTTGGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 715
Oy 307 TCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 366
Db 716 TCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 775
Oy 367 GGTGGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 426
Db 776 GGTGGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 835
Oy 427 GCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 486
Db 836 GCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 895
Oy 487 CTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 546
Db 896 CTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 955
Oy 547 ATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 606
Db 956 ATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1015
Oy 607 GTTGGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 666
Db 1016 GTTGGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1075
Oy 667 GTTGGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 726
Db 1076 GTTGGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1135
Oy 727 CCGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 786
Db 1136 CCGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1195
Oy 787 ATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 846
Db 1196 ATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1255
Oy 847 CAAGTGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 906
Db 1256 CAAGTGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1315
Oy 907 AACACCCCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 966
Db 1316 AACACCCCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1375
Oy 967 GACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1026
Db 1376 GACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1435
Oy 1027 GACCCGGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1086
Db 1436 GACCCGGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1495
Oy 1087 GCTCATGAACTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1146
Db 1496 GCTCATGAACTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1555
Oy 1147 TTGAATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1206
Db 1556 TTGAATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1615
Oy 1207 CCTGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1266
```

```
Db 1616 CCTGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1675
Oy 1267 TATGGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1326
Db 1676 TATGGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1735
Oy 1327 TACAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1358
Db 1736 GSCAAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1767

RESULT 2
US-10-176-758-351
; Sequence 351, Application US/10176758
; Publication No. US2003008353A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C104
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 351
; LENGTH: 4407
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-758-351

Query Match 96.9%; Score 1316.8; DB 9; Length 4407;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Oy 7 GCCATGTCCTCCAGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 66
Db 416 GCCATGTCCTCCAGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 475
Oy 67 GCCCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 126
Db 476 GCCCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 535
Oy 127 CTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 186
Db 536 CTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 595
Oy 187 GAGATGCTGTTTCCAGAGAAAGTCAAGGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 246
Db 596 GAGATGCTGTTTCCAGAGAAAGTCAAGGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 655
Oy 247 AGGCTGTTGTCGCGCTTGGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 306
Db 656 AGGCTGTTGTCGCGCTTGGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 715
Oy 307 TCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 366
Db 716 TCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 775
Oy 367 GGTGGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 426
Db 776 GGTGGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 835
Oy 427 GCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 486
```

```

Db      836  GCACTCTGCACTGGAGTGGGGAGCCCTGTTAGCCGTTACATATACGGGGGCTGAA 895
Oy      487  CTCACCTCCAGCCCTGAGAGGAGCACCCTAACTGCTGGGGGAGCTGGGGCTAC 546
Db      896  CTCACCTCCAGCCCTGAGAGGAGCACCCTAACTGCTGGGGGAGCTGGGGCTAC 955
Oy      547  ATCTACGCGCGAAGAGTCTCCAGCGGTCAAGGTCCCATGTGCACAGTCAAGGCTCT 606
Db      956  ATCTACGCGCGAAGAGTCTCCAGCGGTCAAGGTCCCATGTGCACAGTCAAGGCTCT 1015
Oy      607  CTTGGAAGCCCCAGCCCCAGCCCCAGAGAGCCCAAGCCCTTTCCTCACTAGTAATTT 666
Db      1016  CTTGGAAGCCCCAGCCCCAGCCCCAGAGAGCCCAAGCCCTTTCCTCACTAGTAATTT 1075
Oy      667  GTGGAGACACTGTGTGGTGGAGATGACAAAGTGGCCGATTCACAGGTGGGGGCTAAG 726
Db      1076  GTGGAGACACTGTGTGGTGGAGATGACAAAGTGGCCGATTCACAGGTGGGGGCTAAG 1135
Oy      727  CGCTACCTGCTAACAGTATGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 786
Db      1136  CGCTACCTGCTAACAGTATGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1195
Oy      787  AATCTGTACACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 846
Db      1196  AATCTGTACACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1255
Oy      847  CAAGTGGGGGCGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 906
Db      1256  CAAGTGGGGGCGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1315
Oy      907  AACACCCCTGAGAGACTGTGGAGCCCTGACACATTTGACACAGCCATTTCTTTAAC 966
Db      1316  AACACCCCTGAGAGACTGTGGAGCCCTGACACATTTGACACAGCCATTTCTTTAAC 1375
Oy      967  GACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1026
Db      1376  GACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1435
Oy      1027  GACCCGCTCGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1086
Db      1436  GACCCGCTCGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1495
Oy      1087  GCTCATGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1146
Db      1496  GCTCATGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1555
Oy      1147  TTGATGTGGCTTTGAGACACTCTCTGCAATGTCTCAATGTCTCAATGTCTCAATGTCT 1206
Db      1556  TTGATGTGGCTTTGAGACACTCTCTGCAATGTCTCAATGTCTCAATGTCTCAATGTCT 1615
Oy      1207  CCTGAGAGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1266
Db      1616  CCTGAGAGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1675
Oy      1267  TATGGGCACTGTCTTGTAGACAAACCAAGAGGCTCCATTTGATCTCTGTGTGTGTGT 1326
Db      1676  TATGGGCACTGTCTTGTAGACAAACCAAGAGGCTCCATTTGATCTCTGTGTGTGTGT 1735
Oy      1327  TACAAGGAGAGATGACAAAGGGGTAGGTGCA 1358
Db      1736  GGCAGAGACTATGATGTGACCGCCAGTGC 1767

```

```

: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Pan, James
: APPLICANT: Smith, Victoria
: APPLICANT: Watanabe, Colin K.
: APPLICANT: Wood, William I.
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P343081C50
: CURRENT APPLICATION NUMBER: US/10/175,737
: PRIOR APPLICATION REMOVED - See file wrapper or Palm
: NUMBER OF SEQ ID NOS: 612
: SEQ ID NO 351
: LENGTH: 4407
: TYPE: DNA
: ORGANISM: Homo Sapien
: US-10-175-737-351

Query Match      96.9%: Score 1316.8: DB 9: Length 4407:
Best Local Similarity 98.4%: Pred. No. 0:
Matches 1330: Conservative 0: Mismatches 22: Indels 0: Gaps 0:

Oy      7  GCCATGTCACAGACAGGCTGCATCCCGGAGGGGCTTGGCAGGGCGCTGTGGGA 66
Db      416  GCCATGTCACAGACAGGCTGCATCCCGGAGGGGCTTGGCAGGGCGCTGTGGGA 475
Oy      67  GCGCAACCCCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 126
Db      476  GCGCAACCCCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 535
Oy      127  CTACTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 186
Db      536  CTACTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 595
Oy      187  GAGATCGTGTTCAGAGAGCTCAAGCGACGCTGTGCTGTGCTGTGCGGACCCCTGCC 246
Db      596  GAGATCGTGTTCAGAGAGCTCAAGCGACGCTGTGCTGTGCTGTGCGGACCCCTGCC 655
Oy      247  AGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 306
Db      656  AGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 715
Oy      307  TCCGCTGTGACAGTGTGAGGGGCTGACAGTGTGCTGTGAGGAGGCGCTGTGCTGTG 366
Db      716  TCCGCTGTGACAGTGTGAGGGGCTGACAGTGTGCTGTGAGGAGGCGCTGTGCTGTG 775
Oy      367  GGTGAGGAGAGGCTGTGACACTTACCTGTGACCTATCAATGTGAGATCGGAGTGGTG 426
Db      776  GGTGAGGAGAGGCTGTGACACTTACCTGTGACCTATCAATGTGAGATCGGAGTGGTG 835
Oy      427  GCATCTGTGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 486
Db      836  GCATCTGTGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 895
Oy      487  CTCACCTCCAGCCCTGAGAGGAGCACCCTAACTGCTGGGGGAGCTGGGGCTAC 546
Db      896  CTCACCTCCAGCCCTGAGAGGAGCACCCTAACTGCTGGGGGAGCTGGGGCTAC 955
Oy      547  ATCTACGCGCGAAGAGTCTCCAGCGGTCAAGGTCCCATGTGCACAGTCAAGGCTCT 606
Db      956  ATCTACGCGCGAAGAGTCTCCAGCGGTCAAGGTCCCATGTGCACAGTCAAGGCTCT 1015
Oy      607  CTTGGAAGCCCCAGCCCCAGCCCCAGAGAGCCCAAGCCCTTTCCTCACTAGTAATTT 666
Db      1016  CTTGGAAGCCCCAGCCCCAGCCCCAGAGAGCCCAAGCCCTTTCCTCACTAGTAATTT 1075
Oy      667  GTGGAGACACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 726
Db      1076  GTGGAGACACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1135
Oy      727  CGCTACCTGCTAACAGTATGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 786

```

## RESULT 3

```

US-10-175-737-351
: Sequence 351, Application US/10175737
: Publication No. US20030013153A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Chen, Jian
: APPLICANT: Desnoyers, Luc
: APPLICANT: Goddard, Audrey

```

Db	1136	CGCTACTGCTAAACAGTATGAGGACGACGACCAAGCCTTTACAGACCCAAAGCATCCGC	11955
Qy	787	AATCCCTGCAGCTTGGTGGTGAAGCTGGGCTAATCTCGGGTCAAGCGAGAGAGGGGCC	846
Db	1196	AATCCCTGCAGCTTGGTGGTGAAGCTCGGGCTAATCTCGGGTCAAGCGAGAGAGGGGCC	12555
Qy	847	CAAGTGGGGCCCAAGTGTGTGCGCAGACCCTGCGAGCTTTGTGCGCGAGAGCGGGGCTC	906
Db	1256	CAAGTGGGGCCCAAGTGTGTGCGCAGACCCTGCGAGCTTTGTGCGCGAGAGCGGGGCTC	13155
Qy	907	AACAACCCCTGAGAGCTCGAGACCTTGACCACTTTGACACAGCCATTCTGTTAACCCGTAG	966
Db	1316	AACAACCCCTGAGAGCTCGAGACCTTGACCACTTTGACACAGCCATTCTGTTAACCCGTAG	13755
Qy	967	GACCTGTGTGGAGTCTCCACTTTGCGACAGCGTGGGTATGTGGTGAATGTGGGACCGTGTGT	10266
Db	1376	GACCTGTGTGGAGTCTCCACTTTGCGACAGCGTGGGTATGTGGTGAATGTGGGACCGTGTGT	14355
Qy	1027	GACCCGGGCTCGGAGACTGTGCACTTTGTGGAGAGATGATGGGGCTCCAGTCAAGCCTTCACTGCT	10866
Db	1436	GACCCGGGCTCGGAGACTGTGCACTTTGTGGAGAGATGATGGGGCTCCAGTCAAGCCTTCACTGCT	14855
Qy	1087	GCTCATGAACTGGGTCTATGTCTTCAACATGTCTCATGACAACTCCAAAGCATGCACTAGT	11466
Db	1496	GCTCATGAACTGGGTCTATGTCTTCAACATGTCTCATGACAACTCCAAAGCATGCACTAGT	15555
Qy	1147	TTGAAATGGGCTTTGAGCACACTCTCGCCCATGTGATGAGGCCCTGTGATGGACTATGTGAT	12066
Db	1556	TTGAAATGGGCTTTGAGCACACTCTCGCCCATGTGATGAGGCCCTGTGATGGACTATGTGAT	16155
Qy	1207	CCTGAGAGGCCCTGTGTGCTCCCTCGAGTGCCTGCGCTTATCACTGACTTCCGTGACATGGC	12666
Db	1616	CCTGAGAGGCCCTGTGTGCTCCCTCGAGTGCCTGCGCTTATCACTGACTTCCGTGACATGGC	16755
Qy	1267	TATGGGCACTGTCTCTTAGACAACAACAGAGAGCTCCATTGCACTCTGCTGTGACTGGGAGAC	13366
Db	1676	TATGGGCACTGTGTCTTTAGACAACAACAGAGAGCTCCATTGCACTCTGCTGTGACTGGGAGAC	17355
Qy	1327	TACAAGGACGACATGTACACAAGGGGTAGGTGCA	1358
Db	1736	GGCAAGGACATGTATGTGTGACCCCGCAGTGGCA	1767

```

RESULT 4
US-10-173-706-351
; Sequence 351, Application US/10173706
; Publication No. US20030022293A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C7
; CURRENT APPLICATION NUMBER: US/10/173.706
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 351
; LENGTH: 4407
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-173-706-351

```

Query Match Similarity 96.9%: Score 1316.8: DB 9: Length 4407:									
Best Local Similarity 98.4%: Pred No. 0:									
Matches 1330: Conservative 0: Mismatches 22: Indels 0: Gaps 0									
QY	7	GCATGTCCTCCAGACAGAGCTGCATGCCGAGGAGGGGCTTGSCAGAGGCGCTGCCTGTGGGA	66						
Db	416	GCATGTCCCAACAGAGCTGCATCCCGGAGGGGCTTGSCAGAGGCGCTGCCTGTGGGA	475						
QY	67	GGCCAAACCTGCTCCTGCTCCCATTTGGCGGCTCCTGGCTGGTGTGGGCTTTTCG	126						
Db	476	GGCCAAACCTGCTCCTGCTCCTCATTTGGCGGCTCCTGGCTGGTGTGGGCTTTTCG	535						
QY	127	CTACTGCTGGCCTTCTCTCCTGGCCCTAGCCGCGGTGGCCAGCCCCCTCCCGGAGAG	186						
Db	536	CTACTGCTGGCCTTCTCCTGGCCCTAGCCGCGGTGGCCAGCCCCCTCCCGGAGAG	595						
QY	187	GAGATCGTGTTCACAGAAAGCTCAACGGCAGCGTCTCTGCTGGTGGGACCCCTGGC	246						
Db	596	GAGATCGTGTTCACAGAAAGCTCAACGGCAGCGTCTCTGCTGGTGGGACCCCTGGC	655						
QY	247	AGGCTTTGTGGCGCTTTGGAGGGCTTTGGGGAGAGCCTCTCTATAGCTGAGCAGGAG	306						
Db	656	AGGCTTTGTGGCGCTTTGGAGGGCTTTGGGGAGAGCCTCTCTATAGCTGAGCAGGAG	715						
QY	307	TCCGGTGTGAGGTCGAGGGGCTGACAGTGCAGTACCTGGGCCAGCGGCTGAGCTGTG	366						
Db	716	TCCGGTGTGAGGTCGAGGGGCTGACAGTGCAGTACCTGGGCCAGCGGCTGAGCTGTG	775						
QY	367	GGTGGAGCAGAGCCTGGCACACTACCTGAGTGGCACCATTAAATGGAGATCCGAGATGGG	426						
Db	776	GGTGGAGCAGAGCCTGGCACACTACCTGAGTGGCACCATTAAATGGAGATCCGAGATGGG	835						
QY	427	GCATCTCTCAGTGGGATGGGGAGCCCTGTTAGGGCTTTACAAATATGGGGGCTGAA	486						
Db	836	GCATCTCTCAGTGGGATGGGGAGCCCTGTTAGGGCTTTACAAATATGGGGGCTGAA	895						
QY	487	CTCCACCTCCAGCCCTGGAGAGGAGGACCCTTAATCTGTGGGGGACCTGGGCTCAC	546						
Db	896	CTCCACCTCCAGCCCTGGAGAGGAGGACCCTTAATCTGTGGGGGACCTGGGCTCAC	955						
QY	547	ATCTTACGCCGGAGAGATCCTGCGACAGCGGTCAAGGTGCCATGTGCAGCTACGCTCT	606						
Db	956	ATCTTACGCCGGAGAGATCCTGCGACAGCGGTCAAGGTGCCATGTGCAGAGCTCTCT	1015						
QY	607	CTTGGAGAGCCCAAGCCCGACAGCCCGCAAGAGCCAGCGCTTTGCTTCACTGAGTACATTT	666						
Db	1016	CTTGGAGAGCCCAAGCCCGACAGCCCGCAAGAGCCAGCGCTTTGCTTCACTGAGTACATTT	1075						
QY	667	GTGGAGACACTGGTGTGGCAGATGACAGATGAGGCGGCATTCCACGGGTGCGGGGCTAAAG	726						
Db	1076	GTGGAGACACTGGTGTGGCAGATGACAGATGAGGCGGCATTCCACGGGTGCGGGGCTAAAG	1135						
QY	727	CGCTACCTGCTAAACATGATGGCAGCAGCAGCCAGGCTTCAAGCAGCCAAAGCATCGC	786						
Db	1136	CGCTACCTGCTAAACATGATGGCAGCAGCAGCCAGGCTTCAAGCAGCCAAAGCATCGC	1195						
QY	787	AATCTCTGACGTTGGTGGTGACTCGGCTAAGTATCTCTGGGTGACAGCGAGAGGGGCC	846						
Db	1196	AATCTCTGACGTTGGTGGTGACTCGGCTAAGTATCTCTGGGTGACAGCGAGAGGGGCC	1255						
QY	847	CAAGTGGGAGCAGTCTCTCCAGACAGCCCTGCGCAGCTTCTGTGCGCTGGCAGCGGAGCTC	906						
Db	1256	CAAGTGGGAGCAGTCTCTCCAGACAGCCCTGCGCAGCTTCTGTGCGCTGGCAGCGGAGCTC	1315						
QY	907	AACACCCCTGAGGACTCGGAGCCCTGACACACTTTCAGACAGCCATTCCTGTTTACCCTCAG	966						
Db	1316	AACACCCCTGAGGACTCGGAGCCCTGACACACTTTCAGACAGCCATTCCTGTTTACCCTCAG	1375						
QY	967	GACCTGTGGAGTCTCCACTTGGCAGACAGCGTGGGATATGGGATGTGGGACCCGCTCT	1026						
Db	1376	GACCTGTGGAGTCTCCACTTGGCAGACAGCGTGGGATATGGGATGTGGGACCCGCTCT	1435						
QY	1027	GACCGGCTCGGAGCTGTCCATTGTGGAGGATGATGGGCTCCAGTCAGCCTTCACTGCT	1086						

|||||  
Db 1436 GACCCGCTCGAGCTGTGATGAGATGATGGGCTCCAGTCCAGCTTCACTGCT 1495  
Qy 1087 GCTATGAACTGGGTCATGCTTCAACATGCTCATGCACTCAACCTCAATGCT 1146  
Db 1496 GCTCATGAACTGGGTCATGCTTCAACATGCTCATGCACTCAACCTCAATGCT 1555  
Qy 1147 TTGAATGGGCTTTGAGACCTCTCGCATGTCATGAGCCCTGATGATGCTATGAT 1206  
Db 1556 TTGAATGGGCTTTGAGACCTCTCGCATGTCATGAGCCCTGATGATGCTATGAT 1615  
Qy 1207 CCTGAGAGCCCTGCTGCTCCCTGCAAGTCCCGCTTCACTGACTTCTGAGCAATGGC 1266  
Db 1616 CCTGAGAGCCCTGCTGCTCCCTGCAAGTCCCGCTTCACTGACTTCTGAGCAATGGC 1675  
Qy 1267 TATGGGACCTGCTCTTGTAGCAACCAAGAGCTCCATGTCATGCTGCTGATGGGAC 1326  
Db 1676 TATGGGACCTGCTCTTGTAGCAACCAAGAGCTCCATGTCATGCTGCTGATGGCT 1735  
Qy 1327 TACAAGACGACGATGACACAGGGGTAGTGA 1358  
Db 1736 GGCAAGGACTATGATGCTGACCGGCAATGCA 1767

RESULT 5  
US-10-175-738-351  
: Sequence 351, Application US/10175738  
: Publication No. US20030022294A1  
: GENERAL INFORMATION:  
: APPLICANT: Baker, Kevin P.  
: APPLICANT: Chen, Jlan  
: APPLICANT: Desnoyers, Luc  
: APPLICANT: Goddard, Audrey  
: APPLICANT: Godowski, Paul J.  
: APPLICANT: Gurney, Austin L.  
: APPLICANT: Pan, James  
: APPLICANT: Smith, Victoria  
: APPLICANT: Watanabe, Colin K.  
: APPLICANT: Wood, William I.  
: APPLICANT: Zhang, Zemin  
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
: FILE REFERENCE: P3430R1C45  
: CURRENT APPLICATION NUMBER: US/10/175,738  
: PRIORITY FILING DATE: 2002-06-19  
: PRIOR APPLICATION REMOVED - See File Wrapper or Palm  
: NUMBER OF SEQ ID NOS: 612  
: SEQ ID NO 351  
: LENGTH: 4407  
: TYPE: DNA  
: ORGANISM: Homo Sapien  
US-10-175-738-351

Query Match 96.9%; Score 1316.8; DB 9; Length 4407;  
Best Local Similarity 98.4%; Pred. No. 0;  
Matches 1330; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 7 GCCATGTCCTCCAGACAGGCTCGCATCCCGAGAGGCTTGGAGGCGGTGGCTGGGGA 66  
Db 416 GCCATGTCCTCCAGACAGGCTCGCATCCCGAGAGGCTTGGAGGCGGTGGCTGGGGA 475  
Qy 67 GCCCAACCTGCTCTCTCCCTCCCATTTGCGGCTCTCTGCTGCTGCTGCTGCTG 126  
Db 476 GCCCAACCTGCTCTCTCCCTCCCATTTGCGGCTCTCTGCTGCTGCTGCTGCTG 535  
Qy 127 CTACTGCTGCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 186  
Db 536 CTACTGCTGCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 595  
Qy 187 GAGATCGTGTTCAGAGAGCTCAAGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 246  
Db 596 GAGATCGTGTTCAGAGAGCTCAAGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 655

Qy 247 AGCTGTGTGCTCCGCTTTCAGAGCCCTTTGGGAGAGCGTGTACTAGACTGAGCAGAGC 306  
Db 656 AGCTGTGTGCTCCGCTTTCAGAGCCCTTTGGGAGAGCGTGTACTAGACTGAGCAGAGC 715  
Qy 307 TCCGCTGTGCAAGTTCAGAGGCTGCACTGCACTGCTGGCCAGAGGCTGAGCTGCTG 366  
Db 716 TCCGCTGTGCAAGTTCAGAGGCTGCACTGCACTGCTGGCCAGAGGCTGAGCTGCTG 775  
Qy 367 GGTGGAGCAGAGCCCTGAGCACTGCACTGCTGCACTGCACTGCACTGCACTGCACTG 426  
Db 776 GGTGGAGCAGAGCCCTGAGCACTGCACTGCTGCACTGCACTGCACTGCACTGCACTG 835  
Qy 427 GCATCTGTCACTGGGATGAGGAGCCCTGTTAGGCTGTTCATATATATGAGGAGGCTGAA 486  
Db 836 GCATCTGTCACTGGGATGAGGAGCCCTGTTAGGCTGTTCATATATATGAGGAGGCTGAA 895  
Qy 487 CTCCACCTTCAGCCCTTGGAGGAGGACCCCTAATCTGCTGGGGACCTGGGGCTCAC 546  
Db 896 CTCCACCTTCAGCCCTTGGAGGAGGACCCCTAATCTGCTGGGGACCTGGGGCTCAC 955  
Qy 547 ATCTACGCTCCGAGAGTCTGCGCAGCGGTCAAGGTCCATGTGCAAGCTCAAGCTCCT 606  
Db 956 ATCTACGCTCCGAGAGTCTGCGCAGCGGTCAAGGTCCATGTGCAAGCTCAAGCTCCT 1015  
Qy 607 CTGGAGACCCCGAGCCCGCAGACCCCGAAGAGCAGCCCTTGTGCTTCACTGATGATTT 666  
Db 1016 CTGGAGACCCCGAGCCCGCAGACCCCGAAGAGCAGCCCTTGTGCTTCACTGATGATTT 1075  
Qy 667 GTGAGACACTGTGTGTGGCAGATGACAAAGATGCCCATTCACAGGTGCGGGCTAAG 726  
Db 1076 GTGAGACACTGTGTGTGGCAGATGACAAAGATGCCCATTCACAGGTGCGGGCTAAG 1135  
Qy 727 CGCTACCTGCTTAACAGTATGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 786  
Db 1136 CGCTACCTGCTTAACAGTATGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1195  
Qy 787 AATCTGTCAAGTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 846  
Db 1196 AATCTGTCAAGTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1255  
Qy 847 CAAGTGGGCGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 906  
Db 1256 CAAGTGGGCGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1315  
Qy 907 AACACCCCTGAGAGCTCGAGCCTGACCATTTGACACAGCATTCTGTTACCCGCTAG 966  
Db 1316 AACACCCCTGAGAGCTCGAGCCTGACCATTTGACACAGCATTCTGTTACCCGCTAG 1375  
Qy 967 GACCTGTGTGAGTCTGCACTTGCAGACGCTGGGTATGCTGATGTGGGACCGCTGT 1026  
Db 1376 GACCTGTGTGAGTCTGCACTTGCAGACGCTGGGTATGCTGATGTGGGACCGCTGT 1435  
Qy 1027 GACCCGGCTCGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1086  
Db 1436 GACCCGGCTCGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1495  
Qy 1087 GCTCAATGAACTGGGTCATGCTTCAACATGCTCAATGAACTGGGTCATGCTGCTGCT 1146  
Db 1496 GCTCAATGAACTGGGTCATGCTTCAACATGCTCAATGAACTGGGTCATGCTGCTGCT 1555  
Qy 1147 TTGAATGGGCTTTGAGACCTCTCGCATGTCATGAGCCCTGATGATGCTATGAT 1206  
Db 1556 TTGAATGGGCTTTGAGACCTCTCGCATGTCATGAGCCCTGATGATGCTATGAT 1615  
Qy 1207 CCTGAGAGCCCTGCTGCTCCCTGCAAGTCCCGCTTCACTGACTTCTGAGCAATGGC 1266  
Db 1616 CCTGAGAGCCCTGCTGCTCCCTGCAAGTCCCGCTTCACTGACTTCTGAGCAATGGC 1675  
Qy 1267 TATGGGACCTGCTCTTGTAGCAACCAAGAGCTCCATGTCATGCTGCTGATGGGAC 1326  
Db 1676 TATGGGACCTGCTCTTGTAGCAACCAAGAGCTCCATGTCATGCTGCTGATGGCT 1735  
Qy 1327 TACAAGACGACGATGACACAGGGGTAGTGA 1358

Db 1736 GGCAAGACTATGATGCTGACCGCCAGTGCCA 1767

## RESULT 6

US-10-175-752-351

; Sequence 351, Application US/10175752

; Publication No. US20030022295A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P343081C60

; CURRENT APPLICATION NUMBER: US/10/175,752

; PRIOR FILING DATE: 2002-06-19

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 351

; LENGTH: 4407

; TYPE: DNA

; ORGANISM: Homo Sapien

US-10-175-752-351

Query Match 96.9%; Score 1316.8; DB 9; Length 4407;

Best Local Similarity 98.4%; Pred. No. 0;

Matches 1330; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 7 GCATGTCTCCACACAGAGCTCCGATCCCGGAGGGCTTGCGAGGGCGCTGTGGGA 66  
 Db 416 GCCATGTCCCAACAGAGCTCCGATCCCGGAGGGCTTGCGAGGGCGCTGTGGGA 475  
 Qy 67 GCCCAACCTGCTCTCTCTCCCATTTGCGGCTCTCTGCTGTGGCTGTGG 126  
 Db 476 GCCCAACCTGCTCTCTCTCTCCCATTTGCGGCTCTCTGCTGTGGCTGTGG 535  
 Qy 127 CTACTGTGCT 186  
 Db 536 CTACTGTGCT 595  
 Qy 187 GAGATGCTGTTCACAGAGAGCTCAAGCGGAGGCTCTCTCTCTCTCTCTCTCT 246  
 Db 596 GAGATGCTGTTCACAGAGAGCTCAAGCGGAGGCTCTCTCTCTCTCTCTCTCT 655  
 Qy 247 AGCGTGTGCT 306  
 Db 656 AGCGTGTGCT 715  
 Qy 307 TCCGGTGTGAG 366  
 Db 716 TCCGGTGTGAG 775  
 Qy 367 GGTGAG 426  
 Db 776 GGTGAG 835  
 Qy 427 GCATGTCTCCACACAGAGCTCCGATCCCGGAGGGCTTGCGAGGGCGCTGAA 486  
 Db 836 GCATGTCTCCACACAGAGCTCCGATCCCGGAGGGCTTGCGAGGGCGCTGAA 895  
 Qy 487 CTCACCTCCACAG 546  
 Db 896 CTCACCTCCACAG 955

Qy 547 ATCTACGCCGGAAGAGAGCTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 606  
 Db 956 ATCTACGCCGGAAGAGAGCTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1015  
 Qy 607 CTTCGAG 666  
 Db 1016 CTTCGAG 1075  
 Qy 667 GTGAG 726  
 Db 1076 GTGAG 1135  
 Qy 727 CCTACCTCTTAACAGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 786  
 Db 1136 CCTACCTCTTAACAGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1195  
 Qy 787 AATCTGAG 846  
 Db 1196 AATCTGAG 1255  
 Qy 847 CAAGTGGGAG 906  
 Db 1256 CAAGTGGGAG 1315  
 Qy 907 AACACCCCTGAG 966  
 Db 1316 AACACCCCTGAG 1375  
 Qy 967 GACCTGTGAG 1026  
 Db 1376 GACCTGTGAG 1435  
 Qy 1027 GACCCGAG 1086  
 Db 1436 GACCCGAG 1495  
 Qy 1087 GCTCATGAG 1146  
 Db 1496 GCTCATGAG 1555  
 Qy 1147 TTGAATGGGCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1206  
 Db 1556 TTGAATGGGCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1615  
 Qy 1207 CCTGAG 1266  
 Db 1616 CCTGAG 1675  
 Qy 1267 TATGGGAG 1326  
 Db 1676 TATGGGAG 1735  
 Qy 1327 TACAAG 1358  
 Db 1736 GGCAAGACTATGATGCTGACCGCCAGTGCCA 1767

## RESULT 7

US-10-176-482-351

; Sequence 351, Application US/10176482

; Publication No. US20030022296A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
;; FILE OF INVENTION: ACIDS ENCODING THE SAME  
;; FILE REFERENCE: P3430R1C70  
;; CURRENT APPLICATION NUMBER: US/10/176.482  
;; PRIORITY FILING DATE: 2002-06-20  
;; PRIOR APPLICATION REMOVED - See File Wrapper or Palm  
;; NUMBER OF SEQ ID NOS: 612  
;; SEQ ID NO 351  
;; LENGTH: 4407  
;; TYPE: DNA  
;; ORGANISM: Homo Sapien  
US-10-176-482-351

Query Match 96.9%; Score 1316.8; DB 9; Length 4407;

Best Local Similarity 98.4%; Pred. No. 0;  
Matches 1330; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 7 GCCATGTCACAGACAGCTCGCATCCCGGAGGGGCTTGGAGGCGCTGGCTGGGGA 66  
DB 416 GCCATGTCACAGACAGCTCGCATCCCGGAGGGGCTTGGAGGCGCTGGCTGGGGA 475  
QY 67 GCCAACCCCTGCTCTGCTCCCATTTGGCGCTCTCTGCTGGTGGTGGCTTCTG 126  
DB 476 GCCAACCCCTGCTCTGCTCCCATTTGGCGCTCTCTGCTGGTGGTGGCTTCTG 535  
QY 127 CTACTGCTGGCTCTCTCTGCTCTGACCCGCTGAGCCGCCCTCCCGGGAGAG 186  
DB 536 CTACTGCTGGCTCTCTCTGCTCTGACCCGCTGAGCCGCCCTCCCGGGAGAG 595  
QY 187 GAGATGCTGTTTCCAGAGAGCTCAACGGGACGCTGCTGCTGGGCAACCCCTGCC 246  
DB 536 GAGATGCTGTTTCCAGAGAGCTCAACGGGACGCTGCTGCTGGGCAACCCCTGCC 655  
QY 247 AGCGTGTGTGCGCTTGGAGGCTTGGGAGAGCGCTGCTGAGAGCTGAGAGAGAC 306  
DB 656 AGCGTGTGTGCGCTTGGAGGCTTGGGAGAGCGCTGCTGAGAGCTGAGAGAGAC 715  
QY 307 TCCGGTGTGACAGTGCAGAGGGGCTGACAGTGCATCTGGGCGGACCGCTGAGCTG 366  
DB 716 TCCGGTGTGACAGTGCAGAGGGGCTGACAGTGCATCTGGGCGGACCGCTGAGCTG 775  
QY 367 GGTGAGCAGACCTTGGACCTGACTGCGCACCATCAATGAGAGATCCGAGTGGTG 426  
DB 776 GGTGAGCAGACCTTGGACCTGACTGCGCACCATCAATGAGAGATCCGAGTGGTG 835  
QY 427 GCATCTGCTGAGTGGAGTGGGAGCGCTGTAAGGCTGTCAATATGCGGGGCTGAA 486  
DB 836 GCATCTGCTGAGTGGAGTGGGAGCGCTGTAAGGCTGTCAATATGCGGGGCTGAA 895  
QY 487 CTCACCTCCACGCCCCCTGGAGGAGGACACCCCTAACTGTGTGGGGGACCTGGGCTCAC 546  
DB 896 CTCACCTCCACGCCCCCTGGAGGAGGACACCCCTAACTGTGTGGGGGACCTGGGCTCAC 955  
QY 547 ATCTACGCGCGGAGAGTCTTCCAGCGGTCAAGTCCCATGTGCAAGCTCAAGCTCTCT 606  
DB 956 ATCTACGCGCGGAGAGTCTTCCAGCGGTCAAGTCCCATGTGCAAGCTCAAGCTCTCT 1015  
QY 607 CTGGAAGCGCGGAGCGGAGCGGAGGAGCGGCTTGTGCTGCTGATGATTT 666  
DB 1016 CTGGAAGCGCGGAGCGGAGCGGAGGAGCGGCTTGTGCTGCTGATGATTT 1075  
QY 667 GTGAGAGACTGGTGTGAGATGACAGATGGCGCATTCACAGGCTGGGGGCTAAAG 726  
DB 1076 GTGAGAGACTGGTGTGAGATGACAGATGGCGCATTCACAGGCTGGGGGCTAAAG 1135  
QY 727 CGCTACTCTTAACAGTATGAGCAGCAGCAGCGCTTCAAGCAGCCCAAGCATCCGC 786  
DB 1136 CGCTACTCTTAACAGTATGAGCAGCAGCAGCGCTTCAAGCAGCCCAAGCATCCGC 1195  
QY 787 AATCCGTAGAGTGTGTGAGTGGCTAGTATGATCTGGGGTCAAGGAGAGGAGGCGCC 846  
DB 1196 AATCCGTAGAGTGTGTGAGTGGCTAGTATGATCTGGGGTCAAGGAGAGGAGGCGCC 1255

QY 847 CAAGTGGGGCCAGTGTGCTGCCAGACCCCTGGGACCTTCTGTGCTGGCAGCGGGGCTC 906  
DB 1256 CAAGTGGGGCCAGTGTGCTGCCAGACCCCTGGGACCTTCTGTGCTGGCAGCGGGGCTC 1315  
QY 907 AACACCCCTGAGAGTGGAGCTGAGCCTTACCACTTTGACACAGCATTCTGTTTACCCTGAG 966  
DB 1316 AACACCCCTGAGAGTGGAGCTGAGCCTTACCACTTTGACACAGCATTCTGTTTACCCTGAG 1375  
QY 967 GACCTGTGGAGTGTGCTGCTGAGCAGACGCTGGGTATGGCTGATGTGGGACCGCTGT 1026  
DB 1376 GACCTGTGTGGAGTGTGCTGCTGAGCAGACGCTGGGTATGGCTGATGTGGGACCGCTGT 1435  
QY 1027 GACCGGCTGGAGCTGTGCTGCTGAGAGATGATGATGGGCTTCCAGCCTTCACTGCT 1086  
DB 1436 GACCGGCTGGAGCTGTGCTGCTGAGAGATGATGATGGGCTTCCAGCCTTCACTGCT 1495  
QY 1087 GCTCATGAACCTGGTGTGCTGCTTCAACATGCTTCAATGACACTCCAAGCATGATCACT 1146  
DB 1496 GCTCATGAACCTGGTGTGCTGCTTCAACATGCTTCAATGACACTCCAAGCATGATCACT 1555  
QY 1147 TTGAATGGGCTTTGAGACACTGCTGCGCATGATGAGCGCGCTGATGGCTGATGGAT 1206  
DB 1556 TTGAATGGGCTTTGAGACACTGCTGCGCATGATGAGCGCGCTGATGGCTGATGGAT 1615  
QY 1207 CCTGAGAGCCCTGCTGCTGCCCTGACAGTGGCCGCTTCACTGATGCTGAGCAATGAC 1266  
DB 1616 CCTGAGAGCCCTGCTGCTGCCCTGACAGTGGCCGCTTCACTGATGCTGAGCAATGAC 1675  
QY 1267 TATGGGCACTGCTCTTACACAAACAGAGCTCCATTCATCTGCTGCTGAGTGGGAC 1326  
DB 1676 TATGGGCACTGCTCTTACACAAACAGAGCTCCATTCATCTGCTGCTGAGTGGGAC 1735  
QY 1327 TACAAGAGACGATGACAAAGGGGTAGTCA 1358  
DB 1736 GCGAAGACTATGATGCTGACCCCACTGCCA 1767

#### RESULT 8

US-10-176-757-351  
; Sequence 351, Application US/10176757  
; Publication No. US20030022297A1

#### GENERAL INFORMATION:

;; APPLICANT: Baker, Kevin P.  
;; APPLICANT: Chen, Jian  
;; APPLICANT: Desnoyers, Luc  
;; APPLICANT: Goddard, Audrey  
;; APPLICANT: Godowski, Paul J.  
;; APPLICANT: Gurney, Austin L.  
;; APPLICANT: Pan, James  
;; APPLICANT: Smith, Victoria  
;; APPLICANT: Watanabe, Colin K.  
;; APPLICANT: Wood, William I.  
;; APPLICANT: Zhang, Zemin  
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
;; FILE OF INVENTION: ACIDS ENCODING THE SAME  
;; FILE REFERENCE: P3430R1C86  
;; CURRENT APPLICATION NUMBER: US/10/176.757  
;; PRIORITY FILING DATE: 2002-06-20  
;; PRIOR APPLICATION REMOVED - See File Wrapper or Palm  
;; NUMBER OF SEQ ID NOS: 612  
;; SEQ ID NO 351  
;; LENGTH: 4407  
;; TYPE: DNA  
;; ORGANISM: Homo Sapien  
US-10-176-757-351

Query Match 96.9%; Score 1316.8; DB 9; Length 4407;

Best Local Similarity 98.4%; Pred. No. 0;  
Matches 1330; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 7 GCCATGTCACAGACAGCTCGCATCCCGGAGGGGCTTGGAGGCGCTGGCTGGGGA 66  
DB 416 GCCATGTCACAGACAGCTCGCATCCCGGAGGGGCTTGGAGGCGCTGGCTGGGGA 475

OY	67	GCACCAACCCGACCCCTGCTCCGACATTGGACCGCTCTCCGGCTGGTGTGGCTCTCTG	126
Db	476	GCCCAACCCCTBCCCTGCTGCTGCCATTTGGCCGCTCTCCGCTGTGTGGCTCTCTG	535
OY	127	CTACTGCTGGGCTCTCTCTCTGCCCCCAGCCCGGGCTGGCCAGCCGCCCTCCGCCGGAGAG	186
Db	536	CTACTGCTGGGCTCTCTCTCTGCCCCCAGCCCGGGCTGGCCAGCCGCCCTCCGCCGGAGAG	595
OY	187	GAGATCCGTTTCCAGAGAAAGCTCAACGGCAGGCTGCTCCGCTCGGGACACCCCTGCC	246
Db	596	GAGATCCGTTTCCAGAGAAAGCTCAACGGCAGGCTGCTCCGCTCGGGACACCCCTGCC	655
OY	247	AGGCTGTTGTCCCGCTTTCAGAGCCCTTTGGGGAAACGCTCTACTAGACTGGACGAGAC	306
Db	656	AGGCTGTTGTGCGCGCTTCGAGAGCCCTTTGGGGAAACGCTCTACTAGAGCTGGACGAGAC	715
OY	307	TCCGGTGTGACAGTTCGAGGGGCTGACAGTGCAGTACCTTGGGCCAGGCGCTGAGCTGCTG	366
Db	716	TCCGGTGTGACAGTTCGAGGGGCTGACAGTGCAGTACCTTGGGCCAGGCGCTGAGCTGCTG	775
OY	367	GCTGAGACAGAGCCTTGGCAGCTACTGACTGGCAGCATCAATGAGAGATCCGAGTCCGGT	426
Db	776	GCTGAGACAGAGCCTTGGCAGCTACTGACTGGCAGCATCAATGAGAGATCCGAGTCCGGT	835
OY	427	GCATCTCTGCACTGGGATGGGGGAGGCCCTGTTAGCGGTTCACATATTCGGGGGGCTGAA	486
Db	836	GCATCTCTGCACTGGGATGGGGGAGGCCCTGTTAGCGGTTCACATATTCGGGGGGCTGAA	895
OY	487	CTCCAGCTCCAGCCCTCGAGAGGAGAGGACCCCTCACTGCTGAGGGAGACCTGGGGCTGAC	546
Db	896	CTCCAGCTCCAGCCCTCGAGAGGAGAGGACCCCTCACTGCTGAGGGAGACCTGGGGCTGAC	955
OY	547	ATCTACGCGCGGAAGTCTCTGCGCAGCGGTCAAGTCCCATGTGCAAGCTCAAGGCTCT	606
Db	956	ATCTACGCGCGGAAGTCTCTGCGCAGCGGTCAAGTCCCATGTGCAAGGCTCTCT	1015
OY	607	CTTGGAAAGCCCCAGCCCCAGACCCCGGAAGGCCAACGCGTTTGCTTCACTGAGTATTT	666
Db	1016	CTTGGAAAGCCCCAGCCCCAGACCCCGGAAGGCCAACGCGTTTGCTTCACTGAGTATTT	1075
OY	667	GTGAGAGACACTGCTGTGTGGCAGATGTGACAAGATGGCCGATTCACAGGTTGCGGGCTTAAG	726
Db	1076	GTGAGAGACACTGCTGTGTGGCAGATGTGACAAGATGGCCGATTCACAGGTTGCGGGCTTAAG	1135
OY	727	CGCTACTGCTTACAGTATGTGACAGCAGCAGCCAGAGGCTTCAAGACCCACCAAGATCCGC	786
Db	1136	CGCTACTGCTTACAGTATGTGACAGCAGCAGCCAGAGGCTTCAAGACCCACCAAGATCCGC	1195
OY	787	AATCTGTGACAGCTTGGGTGTGACCTCGGCTATGATCTGAGGGGTCAAGCGCAGAGAGGGGCC	846
Db	1196	AATCTGTGACAGCTTGGGTGTGACCTCGGCTATGATCTGAGGGGTCAAGCGCAGAGAGGGGCC	1255
OY	847	CAAGTGGGGGCCAGTGTCTGCCAGAGACCTTGGCAGCTTCTGTGCTGGCAGCGGGGCCCTC	906
Db	1256	CAAGTGGGGGCCAGTGTCTGCCAGAGACCTTGGCAGCTTCTGTGCTGGCAGCGGGGCCCTC	1315
OY	907	AACACCCCTCAGAGACTGGAGACCTGACACACTTTGACACAGCAATCTGTTACCCCTAG	966
Db	1316	AACACCCCTCAGAGACTGGAGACCTTGGACACACTTTGACACAGCAATCTGTTACCCCTAG	1375
OY	967	GACCTGTGTGAGAGTCTCACTTTGCGACAGCGTGGATGAGTGGGTGAGTGGGACGCTGT	1026
Db	1376	GACCTGTGTGAGAGTCTCACTTTGCGACAGCGTGGATGAGTGGGTGAGTGGGACGCTGT	1435
OY	1027	GACCCGGCTCGGAGCTGTGCCATTGTGAGAGATGATGGCTCCAGCTTCACTGCT	1086
Db	1436	GACCCGGCTCGGAGCTGTGCCATTGTGAGAGATGATGGCTCCAGCTTCACTGCT	1495
OY	1087	GCTCATCAACGGGTCAATGCTTTCACACATGCTCCATGACACCTCAAGCATGCAATCAGT	1146
Db	1496	GCTCATCAACGGGTCAATGCTTTCACACATGCTCCATGACACCTCAAGCATGCAATCAGT	1555

Qy	1207	CGTAGAGAGCCCTGTCGCCCTCCCTGCACTGCCCGGCTTCATCACTGACCTTCCTGGACAAATGGC	1266
Db	1616	CGTAGAGAGCCCTGTCGCCCTCCCTCCACTGACATGGCCGCTTCATCACTGACCTTCCTGGACAAATGGC	1675
Qy	1267	TATGGGCACTGTCTCTTTCAGCAAAACCCAGAGAGCTCCATTGCAATCTGCGCTGTGACTGGGGAC	1326
Db	1676	TATGGGCACTGTCTCTTTCAGCAAAACCCAGAGAGCTCCATTGCAATCTGCGCTGTGACTTTCCT	1735
Qy	1327	TACAGAGAGCAAGCATGACAAGGGGTGGAGTGA	1358
Db	1736	GGCAGAGACTATATGATGCTGACCCGCCCACTGGCCA	1767

RESULT 9  
US-10-17

; Sequence 351, Application US/10176913  
; Publication No. US20030022298A1  
; Publication No. US20030022298A1

```

; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.

```

;; APPLICANT: Chen, Jian  
;; APPLICANT: Desnoyers, Luc

;; APPLICANT: Goddard, Audrey  
;; APPLICANT: Godowski, Paul

;; APPLICANT: Gurney, Austlin  
; APPLICANT: pan, James

APPLICANT: Smith, Victor  
APPLICANT: Watanabe, Col

;; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin

```

; TITLE OF INVENTION: SECRET
; TITLE OF INVENTION: ACILL
;

```

```

; FILE REFERENCE: P3430R1C6
; CURRENT APPLICATION NUMBER

```

```

; CURRENT FILING DATE: 200
; Prior Application removed

```

```

; NUMBER OF SEQ ID NOS: 612
;
; SEQ ID NO 351

```

```

; LENGTH: 4407
; TYPE: DNA

```

ORGANISM: Homo Sapien  
US-10-176-913-351

Query Match	QF
00 10 10 00 00	

Query Match	5
Best Local Similarity	96
Matches 1330: Conserved	

MACINES 1330, CONSEI VALIA

[illegible]

DB 416 GCCAIGTCCAGACAGG

67 GCGCAACCTGCTCTCCTG

Db 476 GCCCAACCTGCCCTCCTC

QY	127	CTACTGCTGGCCTCTCTC

Db 536 CTACTGCTGGCCTCTCTC

QY 187 GAGATCGTGTTCAGAC

Db 596 GAGATCGTGTTCACAGAC

247 AGGCTGTTGTGCCGCTT

Db 656 AGGCTGTTGTGCCGCTTC

QY 307 TCCGGTGTGCAGGTCGAC

Db 716 TCCGGTGCAGGCTCGAGGGCTGACAGTGCAGTACCTGGGCCAGGGCCCTGAGCTG 775







Db	536	CTACTGCTGGGCTCTCTCTCCCGCCCTCAGGCCCGGGCTGGCAGACCCCTCTCCCGGGGAGAG	595
Qy	187	GAGATCGTGTTCAGAGAAAGCTCAAAGGACGTCGTGCTGGCTGGGACACCCCTGCC	246
Db	596	GAGATCTCTGTTCAGAGAAAGCTCAAAGGACGTCGTGCTGGCTGGGACACCCCTGCC	655
Qy	247	AGGCTGTGTGGCGCTTGCGAGGCGCTTGGGGAGAGCGGTACTAGACTGTGGACAGAG	306
Db	656	AGGCTGTGTGGCGCTTGCGAGGCGCTTGGGGAGAGCGGTACTAGACTGTGGACAGAG	715
Qy	307	TCGCGTGTGAGGTCGAGAGGGGCTGAGACGTGCACTACCTGGGCGAGCGCCCTGACCTCG	366
Db	716	TCCGCTGTGACGGTCGAGGGGGCTGACAGGTGCACTACCTGGGCGAGCGCCCTGACCTCG	775
Qy	367	GGTGGACACAGCGCTGGCACTACCTGACTGGCCACCATCAATGGAGATCCGGAGATCCGCTG	426
Db	776	GGTGGACACAGCGCTGGCACTACCTGACTGGCCACCATCAATGGAGATCCGGAGATCCGCTG	835
Qy	427	GCACTCTGTGACTGTGGAGGGGAGCGCTGTTAGCGGTTCACAATATCGGGGGGTGAA	486
Db	836	GCACTCTGTGACTGTGGAGGGGAGCGCTGTTAGCGGTTCACAATATCGGGGGGTGAA	895
Qy	487	CTCCACTCCAGCCCTTGGAGGGAGGCAACCCCTAACTGCTGTGGGGACCTGGGGCTCAC	546
Db	896	CTCCACTCCAGCCCTTGGAGGGAGGCAACCCCTAACTGCTGTGGGGAGCACTGGGGGCTCAC	955
Qy	547	ATTCAGCGCGGAAAGATTCCTGGCAGCGGTGCAAGGTCGCAATGTGCAACAGTCAAGCTCT	606
Db	956	ATTCAGCGCGGAAAGATTCCTGGCAGCGGTGCAAGGTCGCAATGTGCAACAGTCAAGCTCT	1015
Qy	607	CTTGAAGCCCCAGCCCCAGAGCCCGAAGACCAAGCGCTTGTGCTTCACTAGTAGATT	666
Db	1016	CTTGAAGCCCCAGCCCCAGAGCCCGAAGACCAAGCGCTTGTGCTTCACTAGTAGATT	1075
Qy	667	GTGGAGACACTGGGTGGGTGAGATGACAGATGGCGGATTCACAGGTGGGGGCTAAAG	726
Db	1076	GTGGAGACACTGGGTGGGTGAGATGACAGATGGCGGATTCACAGGTGGGGGCTAAAG	1135
Qy	727	CGCTACTGTCTAACAGTGAATGGGAGGACGAGGCAAGGCGCTTTCAAACACCCAAAGATCCG	786
Db	1136	CGCTACTGTCTAACAGTGAATGGGAGGACGAGGCAAGGCGCTTTCAAACACCCAAAGATCCG	1195
Qy	787	AATCTGTCAAGCTTGGTGGTGACTCGGCTTAGTATGATCTGTTGGGTACAGCGGAGGAGGCC	846
Db	1196	AATCTGTCAAGCTTGGTGGTGACTCGGCTTAGTATGATCTGTTGGGTACAGCGGAGGAGGCC	1255
Qy	847	CAAGTGGGAGCCCAAGTGTGTGGCCAGACCCGTGCGAGCTTCTGTGCTCGGCAGCGGGGCTC	906
Db	1256	CAAGTGGGAGCCCAAGTGTGTGGCCAGACCCGTGCGAGCTTCTGTGCTCGGCAGCGGGGCTC	1315
Qy	907	AACACCCCTGAGGACTGCGACCCCTGACCACTTTGACACAGCCATTCTGTTTACCCTGAC	966
Db	1316	AACACCCCTGAGGACTGCGGCGCCTGACCACTTTGACACAGCCATTCTGTTTACCCTGAC	1375
Qy	967	GACCTGTGTGAGTCTCCACTTTGGGACAGCGTGGGTATGGGTGATGGTGGAGCCGTCTGT	1026
Db	1376	GACCTGTGTGAGTCTCCACTTTGGGACAGCGTGGGTATGGGTGATGGTGGAGCCGTCTGT	1435
Qy	1027	GACCGGCTCGAGCTTGCATTTGGTAGAGATGATGGGCTTCAGTCAAGCCCTTCACTGCT	1086
Db	1436	GACCGGCTCGAGCTTGCATTTGGTAGAGATGATGGGCTTCAGTCAAGCCCTTCACTGCT	1495
Qy	1087	GCTATGAACTGGGTCAATGTCTTCAACATGTCTCATGACACTCCAAAGCCATGCATCACT	1146
Db	1496	GCTATGAACTGGGTCAATGTCTTCAACATGTCTCATGACACTCCAAAGCCATGCATCACT	1555
Qy	1147	TTGAAATGGGCTTTGAGCACACTCTGCCCAATGTCAATGAGGCCCTGTATGGCTCATGTGAT	1206
Db	1556	TTGAAATGGGCTTTGAGCACACTCTGCCCAATGTCAATGAGGCCCTGTATGGCTCATGTGAT	1615
Qy	1207	CCTGAGAGCCCTGTGCTCCCTCTCAGTGCGCCGCTTCATCATGACTTCTTGGACATATGAC	1266
Db	1616	CCTGAGAGCCCTGTGCTCCCTCTCAGTGCGCCGCTTCATCATGACTTCTTGGACATATGAC	1675

QY 1267 TATGGGCACTGCTCTTAAACAACAGAGGCTCCATTGCATCTGCTGTGACTGGGAC 1326  
DB 1676 TATGGGCACTGCTCTTAAACAACAGAGGCTCCATTGCATCTGCTGTGACTGGGAC 1735  
QY 1327 TACAAGAGAGGAGTACAGAGGGGTAGTCGA 1358  
DB 1736 GGCAGAGACTATGATGCTGACCGCCAGTGC 1767

RESULT 13  
US-10-174-572-351  
Sequence 351, Application US/10174572  
Publication No. US20030027263A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE OF INVENTION: ACIDS ENCODING THE SAME  
FILE REFERENCE: P3430R1C40  
CURRENT APPLICATION NUMBER: US/10/174,572  
PRIORITY FILING DATE: 2002-06-18  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 351  
LENGTH: 4407  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-174-572-351

Query Match 96.9%; Score 1316.8; DB 9; Length 4407;  
Best Local Similarity 98.4%; Pred. No. 0;  
Matches 1330; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 7 GCCATGTCCCAACAGAGGCTCCGATCCCGGAGGGCTTGCGAGGGCTGTGGGGA 66  
DB 416 GCCATGTCCCAACAGAGGCTCCGATCCCGGAGGGCTTGCGAGGGCTGTGGGGA 475  
QY 67 GCCCAACCTGCTCTCTCTCCCATTTGCGGCTCTCTGCTGGTGGTGGTCTTG 126  
DB 476 GCCCAACCTGCTCTCTCTCCCATTTGCGGCTCTCTGCTGGTGGTGGTCTTG 535  
QY 127 CTACTGCTGGCTCTCTCTCTGCTGACCGCGGCTGGCCAGCCCTCCCGGGAGAG 186  
DB 536 CTACTGCTGGCTCTCTCTCTGCTGACCGCGGCTGGCCAGCCCTCCCGGGAGAG 595  
QY 187 GAGATCGGTTTCCAGAGAGAGTCAAGCGGACGCTGCTGCTGCTGCTGCTGCTG 246  
DB 596 GAGATCGGTTTCCAGAGAGAGTCAAGCGGACGCTGCTGCTGCTGCTGCTGCTG 655  
QY 247 AGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 306  
DB 656 AGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 715  
QY 307 TCCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 366  
DB 716 TCCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 775  
QY 367 GGTGAGAGAGAGGCTGCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 426  
DB 776 GGTGAGAGAGAGGCTGCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 835  
QY 427 GCATCTGCACTGGAGATGGGAGGAGCCCTGTAGGCGTGTTCATATTCGGGGGCTGAA 486

DB 836 GCATCTGCACTGGAGATGGGAGGAGCCCTGTAGGCGTGTTCATATTCGGGGGCTGAA 895  
QY 487 CTCACCTTCAGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 546  
DB 896 CTCACCTTCAGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 955  
QY 547 ATCTAAGGCGGAGAGAGTCTGCGAGCGGCTGCAAGTCTCCATGCAAGCTCAAGCTCT 606  
DB 956 ATCTAAGGCGGAGAGAGTCTGCGAGCGGCTGCAAGTCTCCATGCAAGCTCAAGCTCT 1015  
QY 607 CTGGAAGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 666  
DB 1016 CTGGAAGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1075  
QY 667 GTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 726  
DB 1076 GTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1135  
QY 727 GCGTACCTGCTAAGAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 786  
DB 1136 GCGTACCTGCTAAGAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1195  
QY 787 AATCTGTAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 846  
DB 1196 AATCTGTAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1255  
QY 847 CAAGTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 906  
DB 1256 CAAGTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1315  
QY 907 AACACCCCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 966  
DB 1316 AACACCCCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1375  
QY 967 GACCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1026  
DB 1376 GACCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1435  
QY 1027 GACCGGCTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1086  
DB 1436 GACCGGCTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1495  
QY 1087 GCTATGAAGCTGGTATGATGCTTCAACATGCTGATGACACTCCAGAGCATGATCA 1146  
DB 1496 GCTATGAAGCTGGTATGATGCTTCAACATGCTGATGACACTCCAGAGCATGATCA 1555  
QY 1147 TTGAATGGGCTTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1206  
DB 1556 TTGAATGGGCTTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1615  
QY 1207 CCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1266  
DB 1616 CCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1675  
QY 1267 TATGGGCACTGCTCTTAAACAACAGAGGCTCCATTGCATCTGCTGTGACTGGGAC 1326  
DB 1676 TATGGGCACTGCTCTTAAACAACAGAGGCTCCATTGCATCTGCTGTGACTGGGAC 1735  
QY 1327 TACAAGAGAGGAGTACAGAGGGGTAGTCGA 1358  
DB 1736 GGCAGAGACTATGATGCTGACCGCCAGTGC 1767

RESULT 14  
US-10-174-579-351  
Sequence 351, Application US/10174579  
Publication No. US20030027264A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.

APPLICANT: Gurney,Austin L.  
APPLICANT: Pan,James  
APPLICANT: Smith,Victoria  
APPLICANT: Matanabe,Colin K.  
APPLICANT: Wood,William I.  
APPLICANT: Zhang,Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C31  
CURRENT APPLICATION NUMBER: US/10/174,579  
CURRENT FILING DATE: 2002-06-18  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 351  
LENGTH: 4407  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-174-579-351

Query Match 96.9%; Score 1316.8; DB 9; Length 4407;  
Best Local Similarity 98.4%; Pred. No. 0;  
Matches 1330; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 7 GCCATGTCACAGACAGGCTGACATCCCGGAGGGGCTTGGCAGGGCCCTGCTGTGGGA 66  
DB 416 GCCATGTCACAGACAGGCTGACATCCCGGAGGGGCTTGGCAGGGCCCTGCTGTGGGA 475  
QY 67 GCCCAACCCCTGCTCCGCTCCCATGTCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCG 126  
DB 476 GCCCAACCCCTGCTCCGCTCCCATGTCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCG 535  
QY 127 CTACTGCTGCT 186  
DB 536 CTACTGCTGCT 595  
QY 187 GAGATGCTGTTTCCAGAGAGCTCAAGGCGAGCTCTGCTGCTGCTGCTGCTGCTGCTG 246  
DB 596 GAGATGCTGTTTCCAGAGAGCTCAAGGCGAGCTCTGCTGCTGCTGCTGCTGCTGCTG 655  
QY 247 AGGCTGTTGTCCT 306  
DB 656 AGGCTGTTGTCCT 715  
QY 307 TCCGCTGTCAGTTCAGAGGGCTGACAGTACATCTGAGGCTGAGGCTGAGGCTGAG 366  
DB 716 TCCGCTGTCAGTTCAGAGGGCTGACAGTACATCTGAGGCTGAGGCTGAGGCTGAG 775  
QY 367 GGTGAGCAGAGCTGACCTACCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 426  
DB 776 GGTGAGCAGAGCTGACCTACCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 835  
QY 427 GCATCTGTCAGTTCAGAGGGCTGACAGTACATCTGAGGCTGAGGCTGAGGCTGAG 486  
DB 836 GCATCTGTCAGTTCAGAGGGCTGACAGTACATCTGAGGCTGAGGCTGAGGCTGAG 895  
QY 487 CTCACCTTCAGCCCTGAGAGGAGGACCCCTAATCTGCTGAGGAGGACCTGAGGCTTAC 546  
DB 896 CTCACCTTCAGCCCTGAGAGGAGGACCCCTAATCTGCTGAGGAGGACCTGAGGCTTAC 955  
QY 547 ATCTTACGCGGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 606  
DB 956 ATCTTACGCGGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1015  
QY 607 CTTGAGAGCCCGACCCCGACCCCGAGAGGCAAGGCTTGTCTGCTGCTGCTGCTGCTGCT 666  
DB 1016 CTTGAGAGCCCGACCCCGACCCCGAGAGGCAAGGCTTGTCTGCTGCTGCTGCTGCTGCT 1075  
QY 667 GTGAGAGACTGT 726  
DB 1076 GTGAGAGACTGT 1135  
QY 727 CGCTACCTGCTAACAGT 786

DB 1136 CGCTACCTGCTAACAGT 1195  
QY 787 AATCTGTGAGCTGT 846  
DB 1196 AATCTGTGAGCTGT 1255  
QY 847 CAACTGGGGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 906  
DB 1256 CAACTGGGGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1315  
QY 907 AACACCCCTGAGACTGCTGAGACCTGCTGAGACCTGCTGAGACCTGCTGAGACCTG 966  
DB 1316 AACACCCCTGAGACTGCTGAGACCTGCTGAGACCTGCTGAGACCTGCTGAGACCTG 1375  
QY 967 GACCTGTGTGAGT 1026  
DB 1376 GACCTGTGTGAGT 1435  
QY 1027 GACCCGCTGAGAGCTGCTGAGACCTGCTGAGACCTGCTGAGACCTGCTGAGACCTG 1086  
DB 1436 GACCCGCTGAGAGCTGCTGAGACCTGCTGAGACCTGCTGAGACCTGCTGAGACCTG 1495  
QY 1087 GCTCATGAACTGGGTATGTCTTCAACATGCTCCATGACAACTCCAAAGCATGATAGT 1146  
DB 1496 GCTCATGAACTGGGTATGTCTTCAACATGCTCCATGACAACTCCAAAGCATGATAGT 1555  
QY 1147 TTGAATGGGCTTTGAGACCTGCTGAGACCTGCTGAGACCTGCTGAGACCTGCTGAG 1206  
DB 1556 TTGAATGGGCTTTGAGACCTGCTGAGACCTGCTGAGACCTGCTGAGACCTGCTGAG 1615  
QY 1207 CCTGAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1266  
DB 1616 CCTGAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1675  
QY 1267 TATGGGCACTGTCTTATGAAACCAAGAGCTCATTTGATCTGCTGTACTGGGAGAC 1326  
DB 1676 TATGGGCACTGTCTTATGAAACCAAGAGCTCATTTGATCTGCTGTACTGGGAGAC 1735  
QY 1327 TACAAGGACGAGATGACAAAGGCTGAGTGA 1358  
DB 1736 GGCAGAGACTATGATGCTGACCCGACATGCCA 1767

## RESULT 15

US-10-174-582-351

Sequence 351, Application US/10174582

Publication No. US20030027265A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Chen, Jian

APPLICANT: Desnoyers, Luc

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Pan, James

APPLICANT: Smith, Victoria

APPLICANT: Matanabe, Colin K.

APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE REFERENCE: P3430R1C36

CURRENT APPLICATION NUMBER: US/10/174,582

Prior Application removed - See File Wrapper or Palm

NUMBER OF SEQ ID NOS: 612

SEQ ID NO 351

LENGTH: 4407

TYPE: DNA

ORGANISM: Homo Sapien

US-10-174-582-351

Query Match 96.9%; Score 1316.8; DB 9; Length 4407;

Best Local Similarity 98.4%; Pred. No. 0;  
Matches 1330; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```
OY 7 GCCATGTCACAGACAGGCTGCATCCCGGAGGGGCTGGAGGGGCTGGCTGGCTGGGGA 66
Db 416 GCCATGTCACAGACAGGCTGCATCCCGGAGGGGCTGGAGGGGCTGGCTGGGGA 475
OY 67 GCCCAACCCCTGCTCCCTGCTCCCAATTCGCGCTCTCTGCTGGTGGTGGCTGGTTCG 126
Db 476 GCCCAACCCCTGCTCCCTGCTCCCAATTCGCGCTCTCTGCTGGTGGTGGCTGGTTCG 535
OY 127 CTACTGCTGGGCTCTCTCTGCTCCCTGCTCCAGCCGGCTGGGACAGCCCTCTCCCGGGAGAG 186
Db 536 CTACTGCTGGGCTCTCTCTGCTCCCTGCTCCAGCCGGCTGGGACAGCCCTCTCCCGGGAGAG 595
OY 187 GAGATCGGTTTCAGAGACCTCAAGCGGACGCTGCTGGCTGGGACAGCCCTGCTCC 246
Db 596 GAGATCGGTTTCAGAGACCTCAAGCGGACGCTGCTGGCTGGGACAGCCCTGCTCC 655
OY 247 AGGCTGTTGTCCCGCTTTCAGGCGCTTTGGGAGAGCGCTGCTACTAGACCTGGAGCAGAC 306
Db 656 AGGCTGTTGTCCCGCTTTCAGGCGCTTTGGGAGAGCGCTGCTACTAGACCTGGAGCAGAC 715
OY 307 TCCGCTGCTGCAGCTGCAGGGGCTGACATGCACTGCTGGGCGCAGCGGCTGAGCTGCTG 366
Db 716 TCCGCTGCTGCAGCTGCAGGGGCTGACATGCACTGCTGGGCGCAGCGGCTGAGCTGCTG 775
OY 367 GGTGAGACAGAGGCTGGAGACCTGACCTGCTGACATGCAATGGAGATCCGAGTCCGCTG 426
Db 776 GGTGAGACAGAGGCTGGAGACCTGACCTGCTGACATGCAATGGAGATCCGAGTCCGCTG 835
OY 427 GCATCTCTGCACCTGGAGTGGGGAGCCCTGTTAGCGCTTTACAAATATCGGGGGCTGAA 486
Db 836 GCATCTCTGCACCTGGAGTGGGGAGCCCTGTTAGCGCTTTACAAATATCGGGGGCTGAA 895
OY 487 CTCACCTCTCAGGCGCTTGGAGGGAGCGCCCTAACCTGCTGGGGGACCTGGGGCTGAC 546
Db 896 CTCACCTCTCAGGCGCTTGGAGGGAGCGCCCTAACCTGCTGGGGGACCTGGGGCTGAC 955
OY 547 ATCTACGCGCGGAGAGTCTGCGCAGCGGTCAAGTCCCATGTGCACAGTCAAGGCTCT 606
Db 956 ATCTACGCGCGGAGAGTCTGCGCAGCGGTCAAGTCCCATGTGCACAGTCAAGGCTCT 1015
OY 607 CTTGGAAGCCCGCAGCCGACAGCCCGGAAGAGCAAGCGCTTGGCTTCACTAGTATTT 666
Db 1016 CTTGGAAGCCCGCAGCCGACAGCCCGGAAGAGCAAGCGCTTGGCTTCACTAGTATTT 1075
OY 667 GTGGAGACACTGTGTGGTGGAGATGACAAGATGGCCGATTCCAAGGTTGGGGGCTTAAG 726
Db 1076 GTGGAGACACTGTGTGGTGGAGATGACAAGATGGCCGATTCCAAGGTTGGGGGCTTAAG 1135
OY 727 CGCTACCTGCTACAGTATGATGGAGACAGCAGCAGCCCTTCAAGCAGCCCAAGCATCCG 786
Db 1136 CGCTACCTGCTACAGTATGATGGAGACAGCAGCAGCCCTTCAAGCAGCCCAAGCATCCG 1195
OY 787 AATCTGCTCAGCTTGTGTGTGCTAGTGGCTAGTATCTGTGGGCTCAGCGAGAGGGGCC 846
Db 1196 AATCTGCTCAGCTTGTGTGTGCTAGTGGCTAGTATCTGTGGGCTCAGCGAGAGGGGCC 1255
OY 847 CAAAGTGGGGCCAGTGGTGGCCAGACCCCTGGCAGCTTGTGGCTGGCAGCGGGGCTC 906
Db 1256 CAAAGTGGGGCCAGTGGTGGCCAGACCCCTGGCAGCTTGTGGCTGGCAGCGGGGCTC 1315
OY 907 AACACCCCTGAGACTGAGACCTGAGCAGCTTGTGACACAGCAGCATTGTGTTACCGTCA 966
Db 1316 AACACCCCTGAGACTGAGACCTGAGCAGCTTGTGACACAGCAGCATTGTGTTACCGTCA 1375
OY 967 GACCTGTGTGAGTCTGCATTTGGACACGCTGGTATGCTGATGTGGACAGCTGCTGT 1026
Db 1376 GACCTGTGTGAGTCTGCATTTGGACACGCTGGTATGCTGATGTGGACAGCTGCTGT 1435
OY 1027 GACCGGCTGGAGCTGTGCAATTTGTGAGATGATGGGCTCCAGTCAAGCTTCACTGCT 1086
Db 1436 GACCGGCTGGAGCTGTGCAATTTGTGAGATGATGGGCTCCAGTCAAGCTTCACTGCT 1495
OY 1087 GCTCATGAACCTGGGTCATGTCTTCACATGCTCCATGACACCTCCAGCATGATGATGAT 1146
Db 1496 GCTCATGAACCTGGGTCATGTCTTCACATGCTCCATGACACCTCCAGCATGATGATGAT 1555
OY 1147 TTGAATGGGCTTTGAGCAGCCTTCGCGCATGTGATGGCCCTGTGATGGCTCATGTGAT 1206
Db 1556 TTGAATGGGCTTTGAGCAGCCTTCGCGCATGTGATGGCCCTGTGATGGCTCATGTGAT 1615
OY 1207 CCTGAGAGCCCTGTGCTCCCTGCGAGCTCCCGCTTCACTGATGATGATGATGATGAT 1266
Db 1616 CCTGAGAGCCCTGTGCTCCCTGCGAGCTCCCGCTTCACTGATGATGATGATGATGAT 1675
OY 1267 TATGGGCACTGCTCTTTAGACAAACAGAGGCTCATGATGATGATGATGATGATGAT 1326
Db 1676 TATGGGCACTGCTCTTTAGACAAACAGAGGCTCATGATGATGATGATGATGATGAT 1735
OY 1327 TACAAGACAGATGACAGAGGCTTAGGTGA 1358
Db 1736 GCAAGGACTATGATGTGACCGCCAGTGCCA 1767
```

Search completed: April 26, 2003, 09:49:47  
Job time : 327.363 secs

GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 14:30:17 ; Search time 16.7668 Seconds

(without alignments)  
1217.071 Million cell updates/sec

Title: US-10-050-200-9

Perfect score: 2614  
Sequence: 1 MLCWASLLCAFRPLAAV.....LDLPKQILGGDYKDDDDKG 492

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 segs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2549	97.5	930	1	Q9UNA0 homo sapien
2	2079	79.5	930	1	ATPS5_HUMAN
3	829.5	31.7	967	1	ATPS1_MOUSE
4	823	31.5	968	1	ATPS1_MOUSE
5	810	31.0	967	1	ATPS1_MOUSE
6	755.5	28.9	890	1	ATPS8_HUMAN
7	732.5	28.0	905	1	ATPS8_MOUSE
8	658.5	25.2	837	1	ATPS4_HUMAN
9	646	24.7	1629	1	ATPS9_HUMAN
10	589.5	22.6	630	1	ATPS4_RAT
11	527.5	20.2	997	1	ATPS7_HUMAN
12	500	19.1	1593	1	ATPS2_HUMAN
13	488	18.7	1077	1	ATPS1_MOUSE
14	413	15.8	207	1	ATPS5_HUMAN
15	413	15.8	860	1	ATPS6_HUMAN
16	408.5	15.6	1205	1	ATPS3_HUMAN
17	407	15.6	1205	1	ATPS2_BOVIN
18	387	14.8	1211	1	ATPS2_BOVIN
19	336.5	12.9	920	1	ADP9_MOUSE
20	322.5	12.3	956	1	ADP9_MOUSE
21	313.5	12.0	813	1	ADP3_HUMAN
22	295.5	11.3	245	1	ATPS4_BOVIN
23	290.5	11.1	903	1	ADP2_MOUSE
24	288	11.0	824	1	ADP8_HUMAN
25	278.5	10.7	776	1	ADP8_MACRA
26	276	10.6	775	1	ADP8_HUMAN
27	275.5	10.5	826	1	ADP8_MOUSE
28	271	10.4	769	1	ADP1_HUMAN
29	269.5	10.3	774	1	ADP8_MOUSE
30	266.5	10.2	909	1	ADP12_HUMAN
31	263	10.1	814	1	ADP5_HUMAN
32	262	10.0	816	1	ADP5_RAT
33	253.5	9.7	754	1	ADP7_HUMAN

34	249	9.5	788	1	ADP7_MOUSE	O35227 mus musculus
35	247.5	9.5	413	1	ACIA_AGRAC	Q9PW35 aglystodon
36	244	9.3	797	1	ADP3_MOUSE	Q923W9 mus musculus
37	239	9.1	480	1	DISA_TRIGA	P15503 trimeresurus
38	239	9.1	773	1	ADP1_MOUSE	Q91146 mus musculus
39	236	9.0	776	1	ADP7_MACRA	Q28475 macaca fasc
40	233	8.9	906	1	ADP2_HUMAN	Q9P0K1 homo sapien
41	231	8.8	789	1	ADP7_RAT	Q63180 rattus norv
42	230.5	8.8	571	1	DIS1_BOVJA	P30431 boehrops ja
43	229	8.8	815	1	ADP5_MOUSE	O88839 mus musculus
44	229	8.8	857	1	ADP2_MOUSE	Q91146 mus musculus
45	219	8.4	478	1	HRT6_CROAT	P34182 croatalus at

## ALIGNMENTS

RESULT 1  
ID ATP5\_HUMAN STANDARD: PRT: 930 AA.  
AC Q9UNA0: Q9UKP2:  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE ADPMTS-5 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase  
DE with thrombospondin motifs 5) (ADAM-TS 5) (ADAM-TS5) (Aggrecanase-2)  
DE (ADMP-2) (ADAM-TS 11).  
GN ADPMTS OR ADMP2 OR ADPMTS11.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OC NCBI\_TaxId:9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE: liver;  
RX MEDLINE:99367476; PubMed-10438522;  
RA Abbaszade T., Liu R.-Q., Yang F., Rosenfield S.A., Ross O.H.,  
RA Link J.R., Ellis D.M., Tortorella M.D., Pratta M.A., Hollis J.M.,  
RA Wynn R., Duke J.L., George H.J., Hillman M.C. Jr., Murphy K.,  
RA Wiswall B.H., Copeland R.A., Decicco C.P., Bruckner R., Nagase H.,  
RA Ito Y., Newton R.C., Magolda R.L., Trzaskos J.M., Hollis G.F.,  
RA Arner E.C., Burn T.C.,  
RA "Cloning and characterization of ADPMTS11, an aggrecanase from the  
RA ADPMTS family.";  
RT J. Biol. Chem. 274:23443-23450(1999).  
RL [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE:20289799; PubMed-10830953;  
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,  
RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,  
RA Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,  
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,  
RA Reichwald K., Rump A., Schillabel M., Schudy A., Zimmermann W.,  
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,  
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,  
RA Minoshima S., Shimizu N., Nordliek G., Hornschrader K., Brundt P.,  
RA Schaefer M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecher H.,  
RA Ramer J., Beck A., Klages S., Hennig S., Resselmann L., Dagand E.,  
RA Wehmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,  
RA Lehach H., Reinhardt R., Yaspo M.-L.,  
RA "The DNA sequence of human chromosome 21.";  
RT Nature 405:311-319(2000).  
RN [3]  
RP SEQUENCE OF 413-930 FROM N.A.  
RC TISSUE: fetal brain;  
RX MEDLINE:99395124; PubMed-10464288;  
RA Hunkainen T.L., Hirohata S., Selgin M.F., Apte S.S.,  
RA "ADAM-TS5, ADAM-TS6, and ADAM-TS7, novel members of a new family of  
RT zinc metalloproteases.";  
RL J. Biol. Chem. 274:25555-25563(1999).  
CC -I- FUNCTION: CLEAVES AGGREGAN, A CARTILAGE PROTEOLYCAN, AND MAY BE  
CC INVOLVED IN ITS TURNOVER. MAY PLAY AN IMPORTANT ROLE IN THE  
CC DESTRUCTION OF AGGREGAN IN ARTHRITIC DISEASES. MAY PLAY A ROLE IN

CC PROTEOLYTIC PROCESSING MOSTLY DURING THE PERI-IMPLANTATION PERIOD.  
 CC -1- CATALYTIC ACTIVITY: Cleaves aggrecan at the 392-Glu-1-Ala-393 site.  
 CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED AT LOW LEVEL. PRIMARILY IN PLACENTA BUT ALSO IN OTHER TISSUES, SUCH AS HEART AND BRAIN, AND ALSO CERVIX, UTERUS, BLADDER, ESOPHAGUS, RIB CARTILAGE, CHONDROBLASTOMA, FIBROUS TISSUE AND JOINT CAPSULE FROM AN ARTHRITIC PATIENT.  
 CC -1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.  
 CC -1- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.  
 CC -1- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 2 TSP TYPE-1 DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/announce/or\\_send\\_an\\_email\\_to\\_license@isb-sib.ch](http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF142099: AAD95577.1: -  
 CC EMBL: AP001698: BAA95504.1: -  
 CC EMBL: AP001697: BAA95503.1: -  
 CC EMBL: AF14293: AAF02493.1: -  
 CC HSSP: Q9PW35: 1BUD.  
 CC MEROPS: M12.225: -  
 CC Genew: HGNC:221: ADAMTS5.  
 CC MIM: 605007: -  
 CC InterPro: IPR001762: Disintegrin.  
 CC InterPro: IPR002870: Pep\_M12B\_propep.  
 CC InterPro: IPR001590: Repr01ysin.  
 CC InterPro: IPR000884: TSP1.  
 CC InterPro: IPR000130: Zn\_MTPeptidse.  
 CC Pfam: PF00090: tsep\_1: 2  
 CC Pfam: PF01421: Repr01ysin: 1.  
 CC Pfam: PF01562: Pep\_M12B\_propep: 1.  
 CC SMART: SM00209: TSP1: 2.  
 CC PROSITE: PS00215: ADAM\_MEROP: 1.  
 CC PROSITE: PS00427: DISINTEGRIN\_1: FALSE\_NEG.  
 CC PROSITE: PS00442: ZINC\_PROTEASE: 1.  
 CC PROSITE: PS00922: TSP1: 1.  
 CC Hydrolase: Metalloprotease; zinc; Signal; Glycoprotein; zymogen;  
 CC Repeat: Extracellular matrix.  
 CC SIGNAL: 1 16  
 CC PROPEP: 17 261  
 CC CHAIN: 262 930  
 CC SITE: 209 209  
 CC METAL: 411 410  
 CC ACT\_SITE: 411 411  
 CC METAL: 414 414  
 CC METAL: 420 420  
 CC METAL: 485 485  
 CC DOMAIN: 567 623  
 CC DOMAIN: 624 731  
 CC DOMAIN: 732 874  
 CC DOMAIN: 875 930  
 CC DOMAIN: 37 41  
 CC DOMAIN: 257 261  
 CC CARBOHD: 498 498  
 CC CARBOHD: 728 728  
 CC CARBOHD: 802 802  
 CC CARBOHD: 807 807  
 CC CONFLICT: 138 138  
 CC CONFLICT: 614 614  
 CC CONFLICT: 692 692  
 CC SEQUENCE: 930 AA: 101715 MW: B64281502F8193B CRC64:

Query Match 97.5%; Score 2549; DB 1; Length 930;  
 Best Local Similarity 99.8%; Pred. No. 1.7e-180;  
 Matches 481; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 CC 1 MLGNASLLCAFRPLAAGVPAATPAODKACQPTAAAOPTRRROGEEVOERAEPGH 60  
 CC 1 MLGNASLLCAFRPLAAGVPAATPAODKACQPTAAAOPTRRROGEEVOERAEPGH 60  
 CC 61 PHLAORRRSKGLVONIDLYSGGKVGVLVYAGGRFLDLERDQSVGIACFVPAAGGT 120  
 CC 61 PHLAORRRSKGLVONIDLYSGGKVGVLVYAGGRFLDLERDQSVGIACFVPAAGGT 120  
 CC 121 SAPWRRSHCFRGVYDGSFRLAFAFDLGGDLGDFEAVVHARVTLKPLRGMAEEKGR 180  
 CC 121 SAPWRRSHCFRGVYDGSFRLAFAFDLGGDLGDFEAVVHARVTLKPLRGMAEEKGR 180  
 CC 181 VYDGSARILHYVYREGSFELPPRASCETPASTPEAHENAPAINSPSGRAALASQLLD 240  
 CC 181 VYDGSARILHYVYREGSFELPPRASCETPASTPEAHENAPAINSPSGRAALASQLLD 240  
 CC 241 OSALSPAGSGPOTWRRRRRSISAROVELLVADASARLYRGLOHYLLTLASIANR 300  
 CC 241 OSALSPAGSGPOTWRRRRRSISAROVELLVADASARLYRGLOHYLLTLASIANR 300  
 CC 301 LYSHASIENHILAVVKKVVLGDKDKSLEVSNAATTLKPFCKWOHQHQLGDDEHYD 360  
 CC 301 LYSHASIENHILAVVKKVVLGDKDKSLEVSNAATTLKPFCKWOHQHQLGDDEHYD 360  
 CC 361 AAILFTREDLCGHSCDPLGMADVGTICSPERSCAVIEDDGLAAFTVAHEIGHLGLSH 420  
 CC 361 AAILFTREDLCGHSCDPLGMADVGTICSPERSCAVIEDDGLAAFTVAHEIGHLGLSH 420  
 CC 421 DSKFCEETFGSTEDKRLMSITLTSIDASKPMSKTSATITEFLDGHGNCLLDPRKOI 480  
 CC 421 DSKFCEETFGSTEDKRLMSITLTSIDASKPMSKTSATITEFLDGHGNCLLDPRKOI 480  
 CC 481 LG 482  
 CC 481 LG 482  
 CC 481 LG 482

RESULT 2  
 ATSS\_MOUSE  
 ID ATSS\_MOUSE STANDARD: PRT: 930 AA.  
 AC 09R001:  
 DT 16-OCT-2001 (Rel. 40. Created)  
 DT 16-OCT-2001 (Rel. 40. Last sequence update)  
 DT 15-JUN-2002 (Rel. 41. Last annotation update)  
 DE ADAMTS-5 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 5) (ADAM-TS 5) (ADAM-TS5) (Aggrecanase-2) (ADMP-2) (Implantin).  
 GN ADAMTS5.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99395124; PubMed=10464288;  
 RA Hurskainen T.L., Hirohata S., Seidn M.F., Apte S.S.;  
 RT "ADAM-TS5, ADAM-TS6, and ADAM-TS7, novel members of a new family of zinc metalloproteases".  
 RL J. Biol. Chem. 274:25555-25563(1999).  
 CC -1- FUNCTION: CLEAVES AGGREGAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE INVOLVED IN ITS TURNOVER. MAY PLAY AN IMPORTANT ROLE IN THE DESTRUCTION OF AGGREGAN IN ARTHRITIC DISEASES. MAY PLAY A ROLE IN PROTEOLYTIC PROCESSING MOSTLY DURING THE PERI-IMPLANTATION PERIOD.  
 CC -1- CATALYTIC ACTIVITY: Cleaves aggrecan at the 392-Glu-1-Ala-393 site.  
 CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).



CC -1- DEVELOPMENTAL STAGE: EXPRESSED SPECIFICALLY IN THE PERI-  
 CC IMPLANTATION PERIOD IN EMBRYO AND TROPHOBLAST AND AT LOW OR  
 CC UNDETECTABLE LEVEL THEREAFTER.  
 CC -1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT  
 CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.  
 CC -1- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.  
 CC -1- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 2 TSP TYPE-1 DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@isb-sdb.ch](mailto:license@isb-sdb.ch)).  
 CC -----  
 CC EMBL: AF140673; AAD56356.1; -  
 CC MEROPS: M12.225; -  
 CC MGD: MGI:1346321; Adamts5.  
 CC InterPro: IPR001762; Disintegrin.  
 CC InterPro: IPR002870; Pep\_M12B\_propep.  
 CC InterPro: IPR001590; ReprPolysin.  
 CC InterPro: IPR000884; TSP1.  
 CC InterPro: IPR000130; zn\_MTpeptidse.  
 CC Pfam: PF00090; tsp\_1; 2  
 CC Pfam: PF01421; ReprPolysin; 1.  
 CC Pfam: PF01562; Pep\_M12B\_propep; 1.  
 CC SMART: SM00209; TSP1; 2.  
 CC DR PROSITE: PS50215; ADAM\_MEPRO; 1.  
 CC DR PROSITE: PS50092; TSP1; 1.  
 CC DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 CC DR PROSITE: PS00427; DISINTEGRIN\_L1; FALSE\_NEG.  
 CC KM Hydrolyse: Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;  
 CC Repeat; Extracellular matrix.  
 CC FT SIGNAL 1 21 POTENTIAL.  
 CC FT PROPEP 22 261 POTENTIAL.  
 CC FT CHAIN 262 930 ADAMTS-5.  
 CC FT SITE 209 209 CYSTEINE SWITCH (POTENTIAL).  
 CC FT METIL 410 410 ZINC (CATALYTIC) (BY SIMILARITY).  
 CC FT ACT\_SITE 411 411 BY SIMILARITY.  
 CC FT METIL 414 414 ZINC (CATALYTIC) (BY SIMILARITY).  
 CC FT METIL 420 420 ZINC (CATALYTIC) (BY SIMILARITY).  
 CC FT DOMAIN 485 566 DISINTEGRIN-LINE.  
 CC FT DOMAIN 567 623 TSP TYPE-1 1.  
 CC FT DOMAIN 624 731 CTS-RICH.  
 CC FT DOMAIN 732 874 SPACER.  
 CC FT DOMAIN 875 930 TSP TYPE-1 2.  
 CC FT DOMAIN 41 46 POLY-ALA.  
 CC FT DOMAIN 257 261 POLY-ARG.  
 CC FT CARBOHD 488 498 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHD 728 728 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHD 802 802 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHD 807 807 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC SO SEQUENCE 930 AA; 101780 MM; 84DE84B26170D4DC CRC64;

Query Match 79.5%; Score 2079; DB 1; Length 930;  
 Best Local Similarity 84.0%; Pred. No. 8, 5e-146;  
 Matches 410; Conservative 13; Mismatches 53; Indels 12; Gaps 4;

QY 1 MLGNAS-----LLCAFLPLAAGPAATPAODKAGOPTAAAGAPRRRGEEOVERA 55  
 DB 1 MRLKMSLLLLLLLSASCLSLAADSPPAAPADKTRQROAAAAAEPDQPOSEETIRER 59  
 QY 56 EPPGHPPLAORRRSKGLVONIDOLYSGGKGYLYVYAGGRRLDLERDGSVGIAG-FV 114  
 DB 60 ---GHLOPLAGRRSGGLVNIIDOLYSGGKGYLYVYAGGRRLDLERDVGAGSIV 116  
 QY 115 PAGGTSAPRRHSHCFYRTYVDSRSLAVFLDGLGDSFVAVKARTYLTAKLNGPMA 174  
 DB 117 TAGGGLSASSGHGHCFFYRTYVDSRSLAVFLDGLGDSFVAVKARTYLTAKLNGSMA 176

QY 175 EEKRGVYGDGSRILHYVTRGFSFEALPPRASCETPASTPEAHENPAHNSPGRAL 234  
 DB 177 EVE--RIYDGSRLHYNRCGFSEALPPRASCETPASPSPGSPSVHSRRSRL 234  
 QY 235 ASQLDQSLSPAGSGGPGTWMRRRRRSISRAROVLLVADASMARLYRGLOHYLLTL 294  
 DB 235 AQLLDHSAFSPSGNAGPQTMRRRRRSISRAROVLLVADASMARLYRGLOHYLLTL 294  
 QY 295 ASIARLVSHASIEHRIYAVKVVVLGDKSLVSKNAATTLNFCWQHQLGDD 354  
 DB 295 ASIARLVSHASIEHRIYAVKVVVLGDKSLVSKNAATTLNFCWQHQLGDD 354  
 QY 355 HEHNYDAALFTRFEDLCGHSCDTLGMADVTCISPERSCAVIEDGLHAATVAHEIGH 414  
 DB 355 HEHNYDAALFTRFEDLCGHSCDTLGMADVTCISPERSCAVIEDGLHAATVAHEIGH 414  
 QY 415 LGLSHDSCFCEFTFGSTEDKRLMSILTSIDASKPWSKTSATITTEFLDGHGCLLD 474  
 DB 415 LGLSHDSCFCEFTFGSTEDKRLMSILTSIDASKPWSKTSATITTEFLDGHGCLLD 474  
 QY 475 LPRKOTLG 482  
 DB 475 LPRKOTLG 482

RESULT 3  
 ATSL\_HUMAN STANDARD: PRT: 967 AA.  
 ID ATSL\_HUMAN Q90P80; Q90H83; Q9P2K0; Q9NSJ8;  
 AC Q90H18; Q90P80; Q90H83; Q9P2K0; Q9NSJ8;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase  
 DE with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1) (METH-1).  
 GN ADAMTS1 OR METH1 OR KIAA1346.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Casas C., Pritchard M.A., Estivill X., Arbones M.L.;  
 RT "Cloning, characterization and mapping on human chromosome 21 of the  
 RT orthologue of murine Adamts-1";  
 RL Submitted (Jul-1999) to the EMBL/Genbank/DDJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A., AND FUNCTION.  
 RA Tissue:Heart;  
 RX MEDLINE=99367466; PubMed=10438512;  
 RA Vazquez F., Hastings G., Ortega M.-A., Lane T.F., Oikemus S.,  
 RA Lombardo M., Trueta-Arispe M.L.;  
 RT "METH-1, a human ortholog of ADAMTS-1, and METH-2 are members of a new  
 RT family of proteins with angio-inhibitory activity.";  
 RL J. Biol. Chem. 274:23349-23357(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Tissue:Endothelial cells;  
 RX MEDLINE=20247184; PubMed=10785405;  
 RA Gienke J., Schmitz A.O., Pillarsky C., Hinzmann B., Weiss B.,  
 RA Rosenthal A., Thierach K.H.;  
 RT "Differential gene expression by endothelial cells in distinct  
 RT angiogenic states.";  
 RL Eur. J. Biochem. 267:2820-2830(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Tissue:Brain;  
 RX MEDLINE=20181126; PubMed=10718198;  
 RA Nagase T., Kikuno R., Ishikawa K.-I., Hirose M., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes XVI.  
 RT The complete sequences of 150 new cDNA clones from brain which code  
 RT for large proteins in vitro.";  
 RL DNA Res. 7:65-73(2000).

FT	PROPEP	50	252	BY SIMILARITY.
FT	CHAIN	253	967	ADAMS-1.
FT	SITE	198	198	CYSTEINE SWITCH (POTENTIAL).
FT	METAL	401	401	ZINC (CATALYTIC) (BY SIMILARITY).
FT	ACT_SITE	402	402	ZINC (CATALYTIC).
FT	METAL	405	405	ZINC (CATALYTIC) (BY SIMILARITY).
FT	METAL	411	411	ZINC (CATALYTIC) (BY SIMILARITY).
FT	DOMAIN	476	559	DISINTEGRIN-LIKE.
FT	DOMAIN	560	616	TSP TYPE-1 1.
FT	DOMAIN	617	724	CYS-RICH.
FT	DOMAIN	725	849	SPACER.
FT	DOMAIN	850	908	TSP TYPE-1 2.
FT	DOMAIN	909	967	TSP TYPE-1 3.
FT	DOMAIN	843	846	POLY-LYS.
FT	CARBOHYD	547	547	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	720	720	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	764	764	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	227	227	P -> A (IN REF. 4 AND 5).
FT	CONFLICT	468	468	Q -> H (IN REF. 1).
FT	CONFLICT	561	561	S -> N (IN REF. 1).
SO	SEQUENCE	967 AA;	105383 MM;	C1893893247A1ED1 CRC64;
Query Match 31.7%; Score 829.5; DB 1; Length 967;				
Best Local Similarity 37.8%; Pred. No. 1.5e-53;				
Matches 194; Conservative 54; Mismatches 150; Indels 115; Gaps 13				
QY	26	PAQKAGQEP	-----AAAAGPRRRGGEVGEFAAP-----PGHPPLAQRRSKGL	73
DB	25	PGSRFGVPPTLLLAALLAVSDALGRSEDEDELVELERARQHG---	TTRLRLNAF	81
QY	74	VQNIIDOLYSGGKGVLYVAGGRFLDLERDQSVGIAGV-----	PAGGTSAPWRHR	127
DB	82	DOQDL-----	LELRPSSEFLAPFTLQNVGRKSGSETPLETDL	120
QY	128	SHCFRGTVDSPRSFLAVFDLGGDGFPAVNHARYTLKRLRGPMAREKGRVVGDSGA	187	
DB	121	AHCFTSGVNGDPPSAALSLCEGVRGAFYLLGEATFIQPL--	PAASER-----	167
QY	188	RLHYHTREGFSFEALPPRASCEPTASTPEAH-----	EHAPAHNSPSC	230
DB	168	-----	LATAAPGCKPPAPLQFILRRNRQDVGTCGVVDPEP--	RPTG 209
QY	231	RAALASQLDLOSALSPAGSGSPQTMWR-----	RRRSISRARQVELL	272
DB	210	KA---EIEDEDEEGEGEDEGEPQ--WSQDPALQGVGQPTGTSLRKRFVSSHREVTM	263	
QY	273	LVADASMARLXGRGLQHLVLTASTIANRLYSHASIEHNIIRLAVVWVYVGGDKDSLEYSK	332	
DB	264	LVADQSMAPFHSGSLKHLTLTLFSAARALKRHPSTRNSVSLVYVAILVYHDKQKPEVTS	323	
QY	333	NAATTLLKPKCKWQHONOLGDDHDEHYDAALLFTREDLGGHHSCDTLGMADVGTICSPER	392	
DB	324	NAALLRNFCKWOKOHNPSPDRADAHYTAALLFTRODLGSGTCTPLGMADVGTICDPSR	383	
QY	393	SCAVTEEDGGLAAPFAVEHIGLTLGLSHDSDKFCFCETGSTGSTRDKLMSILTSIDASKPM	452	
DB	384	SCSVTEEDGGLAAPFAHTEHIGLHVFEMPRHDDAKQCSLNGVNDSSHMASSLSNLDHSDPM	443	
QY	453	SKCTSAITTEFLDDGHGNCLLDLPRKQI-LGSD	484	
DB	444	SPCSAYMITTSFLDNGHGECIMDKPQNPIQLPDD	476	
RESULT 4				
ATSI_MOUSE	STANDARD;	PRT;	968 AA.	
AC	39-MAY-2000 (Rel. 39, Created)			
DT	15-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	ADMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with chromobospondin motifs 1) (ADAM-TS 1) (ADAM-TS1).			
DE	ADMTS1			

OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN=129/SVJ;  
 RC MEDLINE=98110583; PubMed=9441751;  
 RA Kuno K., Lizasa H., Ohno S., Matsushima K.;  
 RT "The exon/intron organization and chromosomal mapping of the mouse  
 RT ADAMTS-1 gene encoding an ADAM family protein with TSP motifs.";  
 RL Genomics 46:466-471(1997).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=97150761; PubMed=8995297;  
 RA Kuno K., Kanada N., Nakashima E., Fujiki F., Ichimura F.,  
 RA Matsushima K.;  
 RT "Molecular cloning of a gene encoding a new type of metalloproteinase-  
 RT integrin family protein with thrombospondin motifs as an  
 RT inflammation associated gene.";  
 RL J. Biol. Chem. 272:556-562(1997).  
 RN [3]  
 RN CHARACTERIZATION, AND MUTAGENESIS OF GLU-403.  
 RP MEDLINE=99303657; PubMed=10373500.  
 RA Kuno K., Terashima Y., Matsushima K.;  
 RT "ADAMTS-1 is an active metalloproteinase associated with the  
 RT extracellular matrix.";  
 RL J. Biol. Chem. 274:18821-18826(1999).  
 RN [4]  
 RN FUNCTION.  
 RP MEDLINE=20389568; PubMed=10930576;  
 RA Kuno K., Okada Y., Kawashima H., Nakamura H., Miyasaka M.,  
 RA Ohno H., Matsushima K.;  
 RT "ADAMTS-1 cleaves a cartilage proteoglycan, aggrecan.";  
 RL FEBS Lett. 478:241-245(2000).  
 RN [5]  
 RN FUNCTION, AND INDUCTION.  
 RP MEDLINE=20243757; PubMed=10781075;  
 RA Rober R.L., Russell D.L., Espey L.L., Lydon J.P., O'Malley B.W.,  
 RA Richards J.S.;  
 RT "Progesterone-regulated genes in the ovulation process: ADAMTS-1 and  
 RT cathepsin L proteases.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:4689-4694(2000).  
 CC -1- FUNCTION: CLEAVES AGGREGAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE  
 CC INVOLVED IN ITS TURNOVER. HAS ANGIOGENIC INHIBITOR ACTIVITY (BY  
 CC SIMILARITY). ACTIVE METALLOPROTEASE, WHICH MAY BE ASSOCIATED WITH  
 CC VARIOUS INFLAMMATORY PROCESSES AS WELL AS DEVELOPMENT OF CANCER  
 CC CACHEXIA. MAY PLAY A CRITICAL ROLE IN FOLLICULAR RUPTURE (BY  
 CC SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: CLEAVES AGGREGAN AT THE 1691-GLU-1-LEU-1692  
 CC SITE, WITHIN THE CHONDROITIN SULFATE ATTACHMENT DOMAIN.  
 CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: SECRETED. ASSOCIATED WITH THE EXTRACELLULAR  
 CC MATRIX.  
 CC -1- INDUCTION: INDUCED IN VITRO IN COLON ADENOCARCINOMA CELLS BY  
 CC INTERLEUKIN-1, OR IN VIVO IN KIDNEY AND HEART BY  
 CC LIPOPOLYSACCHARIDE. ALSO INDUCED BY LH STIMULATION IN GRANULOSA  
 CC CELLS OF PROOVULATORY FOLLICLES.  
 CC -1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT  
 CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.  
 CC -1- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.  
 CC -1- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 3 TSP TYPE-1 DOMAINS.  
 CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A  
 CC FRAMESHIFT IN POSITION 7.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC -----  
 CC EMBL: AB001735; BAA24501.1; ALT\_INIT.  
 CC EMBL: D67076; BAA11088.1; ALT\_FRAME.  
 DR MEROPS: M12.222;  
 DR MGD: MGI:109249; Adamts1.  
 DR InterPro: IPR001762; Disintegrin.  
 DR InterPro: IPR002870; Pep\_M12B\_propep.  
 DR InterPro: IPR001590; Reprolysin.  
 DR InterPro: IPR000884; TSP1.  
 DR InterPro: IPR000130; Zn\_Mpeptidase.  
 DR Pfam: PF00090; TSP\_1; 3.  
 DR Pfam: PF01421; Reprolysin; 1.  
 DR Pfam: PF01562; Pep\_M12B\_propep; 1.  
 DR SMART: SM00209; TSP1; 3.  
 DR PROSITE: PS00427; DISINTEGRIN\_1.  
 DR PROSITE: PS50092; TSP1; 3.  
 DR PROSITE: PS50092; TSP1; 3.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 DR Hydrolase: Metalloprotease; zinc; Signal; glycoprotein; zymogen;  
 KW Repeat; Extracellular matrix; Heparin-binding.  
 FT SIGNAL 1 48  
 FT PROPEP 49 253  
 FT CHAIN 254 968  
 FT SITE 206 206  
 FT METAL 402 402  
 FT ACT\_SITE 403 403  
 FT METAL 406 406  
 FT METAL 412 412  
 FT DOMAIN 477 559  
 FT DOMAIN 560 617  
 FT DOMAIN 618 725  
 FT DOMAIN 726 850  
 FT DOMAIN 851 909  
 FT DOMAIN 910 968  
 FT DOMAIN 195 199  
 FT CARBOHYD 548 548  
 FT CARBOHYD 721 721  
 FT CARBOHYD 765 765  
 FT CARBOHYD 783 783  
 FT CARBOHYD 946 946  
 FT MUTGEN 403 403  
 FT CONFLICT 335 335  
 FT CONFLICT 425 425  
 SQ SEQUENCE 968 AA; 105841 MW; 42EBDA5499B6C1 CRC64;  
 Query Match 31.5%; Score 823; DB 1; Length 968;  
 Best Local Similarity 37.3%; Pred. No. 4; Se-53;  
 Matches 192; Conservative 60; Mismatches 151; Indels 112; Gaps 12;  
 OY 1 MLGNASLLCAFRPLPLAVGPAAPADKACQPPYAAAAQPRRQGEV---QERAE 56  
 DB 11  
 DB 35 MLL-----LLASTMLLCARGAHGRPE-----EDELVLPSIERA- 71  
 OY 57 PGCHPHPLAORRRRSKGLVONIDOLVSGGKGVLYVAGRRFLDLERDGSVGIAGF--- 113  
 DB 11  
 DB 72 -PCHDSTTRRLR-----LDAFCQQLHLKIQPPSGFLGAPFTLQ 108  
 OY 114 -VPAGGCTSA-----PARRSHGCFYRGTVDGSPRSVLAVFLDGGDLGDFPAVKARATYLRKL 168  
 DB 11  
 DB 109 TVGRSPSGEAQHLDPDGLAHCFYSGTVNGDPSAALSLICGVRKAFYLGQEEFTIQ- 167  
 OY 169 LQGPMAEERKGRVYGGASRIILHVYTRGEFSFALPPRASCETPASTPAHEHAPAHSNP 228  
 DB 11  
 DB 168 -----APGVATERLAPAVPEERSARPOPH---ILRRRR 198  
 OY 229 SGRALASQLLDQSALSPAGSGPQT---WWR-----RRRSISR 265  
 DB 11  
 DB 199 RSGGAKCQVMDDEL-PTSDSRPESQNRNQPVPDPTPOAGKPSGCSIRKKRFVSS 257  
 OY 266 ARQVELLVADSMALYRGLOHYLLTLASTANRLYSASISNHLRLAVVYVYVLDGD 325  
 DB 11  
 DB 258 PRVETMLVADQSMADPHSGLKHLYLTLESVAAARYKHPISIRNSISLVVYILVIEBO 317

QY 326 KSLVSKNAATTLKNECKWQHONQDDBHEHYDAALLFTREDLCGHSCTDGLGMADVG 385  
 DB 318 KGEVTSNMAALTRNFCNMQKHNSPDRPEHYDAALLFTREDLCGHSCTDGLGMADVG 377  
 QY 386 TITSPERSCAVIDDDGLHAFTVAHEHGLGLSHDSKCECTFGSTEDKRLMSLWS 445  
 DB 378 TVCDPSRSCVIEDDGLQAFTTAHEHGLHYFNPBHDKACASLNGVTGDHSLMASLWS 437  
 QY 446 IDASKPMKSTATITEFLDDHGNCNLLDPRKOI 480  
 DB 438 LDHSQPMSPCSATIMVTSFLDNGHCECLMDKPNPT 472

RESULT 5  
 ATSL\_RAT ID STANDARD: PRT: 967 AA.

AC 09WU01: 09ER11: 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase  
 with thrombospondin motifs 1) (ADAM-TS1).  
 GN ADAMTS1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=Sprague-Dawley; TISSUE=Brain;  
 RA Liu X., Tu Y., Johnstone E.M., Stephenson D.T., Clemens J.A.,  
 RA Little S.P.;  
 RA "Induction of a disintegrin and metalloproteinase with the  
 RA thrombospondin type I motif (ADAMTS).";  
 RT Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 18-967 FROM N.A.  
 RA STRAIN=Sprague-Dawley; TISSUE=Liver;  
 RA MEDLINE=20304099; PubMed=10847486;  
 RA Diamantis I., Luechli M., Hoessli M., Reichen J.;  
 RT "Cloning of the rat ADAMTS-1 gene and its down regulation in  
 RT endothelial cells in cirrhotic rats.";  
 RL Liver 20:165-172(2000).  
 CC -1- FUNCTION: CLEAVES AGGREGAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE  
 CC INVOLVED IN ITS TURNOVER. HAS ANGIOGENIC INHIBITOR ACTIVITY (BY  
 CC SIMILARITY). ACTIVE METALLOPROTEASE, WHICH MAY BE ASSOCIATED WITH  
 CC VARIOUS INFLAMMATORY PROCESSES AS WELL AS DEVELOPMENT OF CANCER  
 CC CACHEXIA. MAY PLAY A CRITICAL ROLE IN FOLLICULAR RUPTURE (BY  
 CC SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: CLEAVES AGGREGAN AT THE 1683-GLU-I-LEU-1684  
 CC SITE, WITHIN THE CHONDROITIN SULFATE ATTACHMENT DOMAIN.  
 CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular  
 CC matrix (By similarity).  
 CC -1- INDUCTION: DOWN-REGULATED IN ENDOTHELIAL CELLS DERIVED FROM  
 CC CIRRHOTIC LIVER.  
 CC -1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT  
 CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.  
 CC -1- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.  
 CC -1- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 3 TSP TYPE-1 DOMAINS.  
 CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: AF149118; AAD34012.1; -  
 DR EMBL: AF304446; AAC29823.1; -  
 DR MEROPS: M12.222; -  
 DR Interpro: IPR001762; Disintegrin.  
 DR Interpro: IPR002870; Rep\_M12B\_propep.  
 DR Interpro: IPR001590; Repolysin.  
 DR Interpro: IPR000884; TSP1.  
 DR Interpro: IPR000130; Zn\_MTPeptidse.  
 DR Pfam: PF00090; tsp\_1; 6.  
 DR Pfam: PF01421; Repolysin; 2.  
 DR Pfam: PF01562; Rep\_M12B\_propep; 2.  
 DR SMART: SM00209; TSP1; 3.  
 DR PROSITE: PS00215; ADAM\_MERO; 1.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 DR PROSITE: PS50092; TSP1; 2.  
 DR PROSITE: PS00427; DISINTEGRIN\_1; FALSE NEG.  
 KW Hydrolase; Metalloproteinase; Zinc; Signal; Glycoprotein; Zymogen;  
 KW Repeat; Extracellular matrix; Heparin-binding.  
 FT SIGNAL 1 54  
 FT PROPEP 55 252  
 FT CHAIN 253 967  
 FT SITE 205 205  
 FT METAL 401 401  
 FT ACT\_SITE 402 402  
 FT METAL 405 405  
 FT METAL 411 411  
 FT DOMAIN 476 558  
 FT DOMAIN 559 615  
 FT DOMAIN 616 724  
 FT DOMAIN 725 857  
 FT DOMAIN 858 907  
 FT DOMAIN 908 967  
 FT DOMAIN 194 198  
 FT CARBOHYD 547 547  
 FT CARBOHYD 720 720  
 FT CARBOHYD 764 764  
 FT CARBOHYD 782 782  
 FT CARBOHYD 945 945  
 FT CARBOHYD 945 945  
 FT CONFLICT 26 31  
 FT CONFLICT 49 49  
 FT CONFLICT 72 72  
 FT CONFLICT 79 79  
 FT CONFLICT 249 249  
 FT CONFLICT 262 265  
 FT CONFLICT 607 607  
 FT CONFLICT 936 936  
 FT CONFLICT 962 962  
 SQ SEQUENCE 967 AA; 105705 MW; F93C864F8DCDB4CF CRC64;

Query Match. 31.0%; Score 810; DB 1; Length 967;  
 Best local Similarity 42.0%; Pred. No. 4, 1e-52;  
 Matches 175; Conservative 53; Mismatches 131; Indels 58; Gaps 7;

QY 93 AGGRFLDLERDGVGNGF-----VPAGGCTSA-----PMRHSRCFYKTVDSGRSLA 144  
 DB 84 AFGQQLHLKLPDPSGLAGFTLQTVGRSPGSEAOHLPTGDLAHCFYSTVNGDPSAA 143  
 QY 145 VPDGLGGLGDFEFAVKARYATLKP-----LIRGMAEE-----KGRVYDGS 187  
 DB 144 ALSLCGVRAFLYLGEEFFIOPAPAVATERLYPAERKESIAIPRRPHILRRRRSGGA 203  
 QY 188 RILHYVTRREGFSFEALPP-----RASCETPASPPEAHENIAPAHNSPGSRAALASQLDQSA 243  
 DB 204 KC-----GVNDEETLPTNSGREGSONTPDQWLRNPPTPGACKPTGPGSI----- 248  
 QY 244 LSPAGSGPQTMRRRRRSISRAROVELLVADASNAKLYGKGLQHYLLTLASIANRLYS 303  
 DB 249 -----RRKRFEVSPRYVETMLVADQSMADFGSLKHYLLTLFVAAAFYK 294  
 QY 304 HASTENHRLAVKVVVLDKDKSLVSKNAATTLKNECKWQHONQDDBHEHYDAAL 363  
 DB 295 HPSIRNISTSLVVKILVIYEOKGPEVTSNAAITLKNPFSWOKHNSPSDROPEHYDTAI 354

QY 364 LEFTREDLCGHSODTLGNADVGTICSPERSCAVIEDDGLHAFTVAHEIGHLGLSHDS 423  
 DB 355 LEFTRODLGSHDCTLGNADVGTICSPERSCAVIEDDGLHAFTVAHEIGHLGLSHDS 414  
 QY 424 KCEETFGSTEDKRLMSSILTSIDASKPWSKCTATITEFLDGHGNCILDPKROI 480  
 DB 415 KHCASFNGVSGDHSIMASMLSLDHSQPSWPCSAVAVTSLFDNGHCECLMDKRPQNT 471

RESULT 6  
 AT88\_HUMAN  
 ID AT88\_HUMAN STANDARD: PRT: 890 AA.  
 AC Q9UPF79; Q9NZS0;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE ADAMTS-8 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase  
 with thrombospondin motifs 8) (ADAM-TS 8) (ADAM-TS8) (METH-2)  
 DE (METH-8).  
 GN ADAMTS8 OR METH2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RX MEDLINE=93367466; PubMed=10438512;  
 RA Vazquez F., Hastings G., Ortega M.-A., Lane T.F., Oikemus S.,  
 RA Lombardi M., Iruela-Arispe M.L.;  
 RT "METH-1, a human ortholog of ADAMTS-1, and METH-2 are members of a new  
 family of proteases with angio-inhibitory activity.";  
 RT J. Biol. Chem. 274:23349-23357(1999).  
 RU [2]  
 RP SEQUENCE OF 195-440 FROM N.A.  
 RX MEDLINE=20079168; PubMed=10610729;  
 RA Georgiadis K.E., Hirohata S., Seldin M.F., Apte S.S.;  
 RT "ADAM-TS8, a novel metalloprotease of the ADAM-TS family located on  
 mouse chromosome 9 and human chromosome 11.";  
 RU Genomics 62:312-315(1999).  
 CC -1- FUNCTION: HAS ANTI-ANGIOGENIC PROPERTIES.  
 CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular  
 matrix (By similarity).  
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN ADULT AND FETAL LUNG. LOWER  
 EXPRESSION IN BRAIN, PLACENTA, HEART, STOMACH AND FETAL BRAIN AND  
 KIDNEY.  
 CC -1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT  
 FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.  
 CC -1- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY  
 SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.  
 CC -1- SIMILARITY: CONTRAINS 1 DISINTEGRIN-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 2 TSP TYPE-1 DOMAINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: AF060153; AADA8081.1; -;  
 DR EMBL: AF175283; AAF25806.1; -;  
 DR HSSP: P34179; 11ng.  
 DR MEMOPS: M12.226; -;  
 DR Genew: HGNC:224; ADAMTS8.  
 DR MIM: 605175; -;  
 DR InterPro: IPR001762; Disintegrin.  
 DR InterPro: IPR002870; Pep\_M12B-propep.  
 DR InterPro: IPR001590; Reprolysin.

DR InterPro: IPR000884; TSP1.  
 DR InterPro: IPR000130; Zn\_MTPeptide.  
 DR Pfam: PF000990; tsp\_1; 2  
 DR Pfam: PF01421; Reprolysin; 1.  
 DR Pfam: PF01562; Pep\_M12B-propep; 1.  
 DR SMART: SM00209; TSP1; 2.  
 DR PROSITE: PS50215; ADAM\_MERO; 1.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 DR PROSITE: PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 DR Hydrolase: Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;  
 KW Repeat: Extracellular matrix; Heparin-binding.  
 FT SIGNAL 1 27  
 FT PROPEP 28 214  
 FT CHAIN 215 890  
 FT METAL 364 364  
 FT ACT\_SITE 365 365  
 FT METAL 368 368  
 FT METAL 374 374  
 FT METAL 439 526  
 FT DOMAIN 527 583  
 FT DOMAIN 584 690  
 FT DOMAIN 691 832  
 FT DOMAIN 833 890  
 FT DOMAIN 202 205  
 FT CARBOHYD 345 345  
 FT CARBOHYD 401 401  
 FT CARBOHYD 466 466  
 FT CARBOHYD 491 491  
 FT CARBOHYD 600 600  
 FT CONFLICT 195 195  
 FT CONFLICT 413 440  
 FT SEQUENCE 890 AA: 96671 MM: 37070E0305739D3 CRC64;  
 SO

Query Match 28.9%; Score 755.5; DB 1; Length 890;  
 Best Local Similarity 36.2%; Pred. No. 3; 9e-48;  
 Matches 187; Conservative 58; Mismatches 165; Indels 107; Gaps 12;

QY 8 LLLCAFRLLAVGPRATPAODKAGOPPTAAANAOORRQGEVGERAEPGHPHPLAOR 67  
 DB 15 LLLLLLLPLARAGPARPAA--GGQ--ASELVVETRLPG----- 49  
 QY 68 RRSKGLVONIDOLYSGCGVGLVYAGRRFLDLDERDSV-----GIAGFPVAGCG 119  
 DB 50 -----SAGELALHLSAFGKGFVRLARPDSFLAPEKIRIGSGCATGC 94  
 QY 120 TSAPMHRSHCEYRGTVDSPSLAVFDLCGLDGFPAVKAHARYTLKPL-LRGPAAEEK 178  
 DB 95 E-----RGLGCFEFGSTVNGEPESLAVSLCRGLSGSFLDGEFTIQPGAGAGSLAQPHR 150  
 QY 179 GRVYGGGSRIL-----HYTRREGSFALPPRASCCEFPASPPEAHENAHANSPGCA 232  
 DB 151 LQKMGAGARPLPRGPEWEVEETGEGORQ-----RGDHQDSEEESEAEAGASEPP--- 203  
 QY 233 ALASQLDQSALSPAGSGSPOTWMRRRSISAROVETLLVADASMAFLYGRGLDHYLL 292  
 DB 204 -----PLIGATG-----RTKRFVSEARVETLLVADASMAFLYGRGLDHYLL 245  
 QY 293 TLASIANRLYSHASIEHNLRLAVKVVVLDGDKSLEVSKNATTLKNCQHOHNOUG 352  
 DB 246 TLMSTVARIYKHPSIKNSINLWVKVLIWEDERKMGREVSNDGGLTLRNCNMORRNGS 305  
 QY 353 DDHEHYDAIILFTREDLCGHS--CETLMADVGTICSPERSCAVIEDDGLHAFTVAHE 411  
 DB 306 DRPHEHYDAIILFTRONFGQEGCLCTGLVADIGTCDNKKSCSVIEDGCLAAHFLAHE 365  
 QY 412 IGHLLGLSHDSKFCETFGSTEDKRLMSSILTSIDASKPWSKCTATITEFLDGHGNC 471  
 DB 366 LGHVLSPRHDDSKPCPTRLFGPMCKHHVMARPLVHLNQTILPMSPCSAAMYLTLLDGHGNC 425  
 QY 472 LLDLPR-----KOLLGGDYK 486  
 DB 111 I

DB 426 LIDAGCALPLPTGLPGRMALYQLDQCKQIFGPDFR 462

RESULT 7

AT8\_MOUSE STANDARD: PRT; 905 AA.

AC P57110. 1 228

BT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE ADAMTS-4 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 8) (ADAM-TS 8) (ADAM-TS8) (METH-2).

GN ADAMTS8.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxId=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=20079168; PubMed=10610729;

RT Georgiadis K.E., Hirohata S., Seldin M.F., Apte S.S.;

RT \*ADAM-TS8, a novel metalloproteinase of the ADAM-TS family located on

RT mouse chromosome 9 and human chromosome 11.";

RT Genomics 62:312-315(1999).

CC -1- FUNCTION: HAS ANTI-ANGIOGENIC PROPERTIES (BY SIMILARITY).

CC -1- CORFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).

CC -1- TISSUE SPECIFICITY: EXPRESSED SPECIFICALLY IN ADULT LUNG AND HEART AND LOW EXPRESSION DURING MOUSE DEVELOPMENT.

CC -1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.

CC -1- PTM: THE PRECURSOR IS CLEAVED BY A FUROS ENDOPEPTIDASE (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.

CC -1- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.

CC -1- SIMILARITY: CONTAINS 2 TSP TYPE-1 DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sdb.ch/announce/> or send an email to [license@isb-sdb.ch](mailto:license@isb-sdb.ch)).

CC -----

DR EMBL: AF175282; AAF25805.1; -

DR HSSP: P34179; 11NG.

DR MEROPS: M12.226; -

DR MGD: MGI:1353468; Adamts8.

DR InterPro: IPR001762; Disintegrin.

DR InterPro: IPR001590; Reptolysin.

DR InterPro: IPR000884; TSP1.

DR InterPro: IPR000130; Zn\_MTPepidase.

DR Pfam: PF00090; tsp.1; 2; MTPepidase.

DR SMART: PF01421; Reptolysin; 1.

DR SMART: SM00209; TSP1; 2.

DR PROSITE: PS00142; ZINC\_PROTEASE; 1.

DR PROSITE: PS50215; ADAM\_MERO; 1.

DR PROSITE: PS50092; TSP1; 1.

DR PROSITE: PS00427; DISINTEGRIN; 1.

DR HydroLase: Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen; Repeat; Extracellular matrix; Heparin-binding.

KW Repeat: Extracellular matrix; Heparin-binding.

FT SIGNAL 1 228

FT PROPEP 29 228

FT CHAIN 229 905

FT METAL 378 378

FT METAL 379 379

FT ACT\_SITE 382 382

FT METAL 388 388

FT DOMAIN 453 541

FT DOMAIN 542 598

FT DOMAIN 599 705

FT CYS-RICH.

FT DOMAIN 706 847

FT CARBOHYD 848 905

FT CARBOHYD 415 415

FT CARBOHYD 480 480

FT CARBOHYD 506 506

FT CARBOHYD 615 615

SO SEQUENCE 905 AA; 98879 MW; 124D4132B3A0CAE CRC64.

Query Match

Best Local Similarity 33.5%; Pred. No. 26-46; Indels 105; Gaps 11;

Matches 177; Conservative 61; Mismatches 185;

4 GNASLLCAFLPLAAGVPAATPAADKACQPTAAAGAPRRRGEVEVERAEP--PGH 60

9 GWPPLLLLLGLP-----PPPLVCGAAGCTGAGASLVVPRRLPS 51

61 PNPPLARRSKGLVONIDOLYSGGKGYLYVYAGGRFLDLERDQSV-----GIAG 112

52 ASALA-----FHLSAFGGFVRLAPASFLAPAFKIERLGG 88

113 FVPAGGTSAPMRHSHCHPFGCTVDSPPSLAVFDLGGDLGDFVAKHARYTLKPLRGP 172

89 SSMAAGGFG-----LRGCFSGTVNGERESLAAMSCVAGMSSFLLAGGEFTIIOGAGD 144

173 -----MAEEKGRVYGGDSARILHYVTRGEGFEPALPPRASCTPASTPEAHERA 222

145 SLDPGRILDRMGPCGRREDPGLAAEVF-----PLPGLMEVEMNGCOGERS 193

223 P-AHNSPSCRAALASQLDQSLSPVAGSGPOTWRRRRRSISRAROVLLVVAASNAK 281

194 DNEEDRKQDEGLKETEDSKRVVPPFGSK-----TRSKRFVSEARFETLLVVAASMA 248

282 LVGRGLQHYLLTASIANLSEYSHASIEHRIAVYKVVVLGDKDSLEYSKNAATLKNF 341

249 FYGTDLQNLILVMSAAIYKHPSIRNSVNLVYKVLVEKERNCEPVSMDGGLTLRNF 308

342 CWOQHONLGDDEHEDDAILFTREDLCGH-HSCDTLGMADVCTGSPERSCAVIEDD 400

309 CSWDRFNPSPDRHEHYDTAILFTROMFCGKGEODCTIGMADVCTGDPKSCSVINKE 368

401 GLHAAPYVAHEHGLLGSNDSKCFEFGSTEDKRLMSILTSIDSKPSPKSKTSATY 460

369 GLQAAVTLAHEHGLHSLMHPHDSKPCVRLFGPMGYHMMAPFTHVNTKLTPSPCSAVYL 428

461 YELLDGHCNCLLDLPR-----KQILGSDYK 486

429 TELLDGHCNCLLDLPR-----KQILGSDYK 476

RESULT 8

AT84\_HUMAN STANDARD: PRT; 837 AA.

AC 075173; G9UNB3;

BT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE ADAMTS-4 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 4) (ADAM-TS 4) (ADAM-TS4) (Aggrecanase 1)

GN ADAMTS4 OR KTA0688.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI\_TaxId=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA TISSUE=Brain;

RA MEDLINE=98403880; PubMed=9734811;

RA Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.,

RA Kotani H., Nomura N., Ohara O.;

RT \*Prediction of the coding sequences of unidentified human genes. X.

RT The complete sequences of 100 new cDNA clones from brain which can

RT code for large proteins in vitro.";

RL DNA Res. 5:169-176(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99286303; PubMed=10356395;  
 RA Tortorella M.D., Burn T.C., Pratta M.A., Abbascade I., Hollis J.M.,  
 RA Liu R.-Q., Rosenfeld S.A., Copeland R.A., Decicco C.P., Wynn R.,  
 RA Rockwell A., Yang F., Duke J.L., Solomon K., George H., Bruckner R.,  
 RA Nagase H., Itoh Y., Ellis D.M., Ross H., Mismail B.H., Murphy K.,  
 RA Hillman M.C. Jr., Hollis G.F., Newton R.C., Magolda R.L.,  
 RA Trzaskos J.M., Arner E.C.;  
 RT "Purification and cloning of aggrecanase-1: a member of the ADAMTS  
 RT family of proteases.";  
 RT Science 284:1664-1666(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Sawaji Y., Nagase H., Saklatvala J., Clark A.R.;  
 RT "ADAMTS-4 genomic locus.";  
 RT Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.  
 RN [4]  
 RP PARTIAL SEQUENCE, AND CHARACTERIZATION.  
 RX MEDLINE=20400518; PubMed=10827174;  
 RA Tortorella M.D., Pratta M.A., Liu R.-Q., Abbascade I., Ross H.,  
 RA Burn T.C., Arner E.C.;  
 RT "The thrombospondin motif of aggrecanase-1 (ADAMTS-4) is critical for  
 RT aggrecan substrate recognition and cleavage.";  
 RT J. Biol. Chem. 275:25791-25797(2000).  
 CC -1- FUNCTION: CLEAVES AGGREGAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE  
 CC INVOLVED IN ITS TURNOVER. MAY PLAY AN IMPORTANT ROLE IN THE  
 CC DESTRUCTION OF AGGREGAN IN ARTHRITIC DISEASES. COULD ALSO BE A ONE  
 CC CRITICAL FACTOR IN THE EXACERBATION OF NEURODEGENERATION IN  
 CC ALZHEIMER'S DISEASE.  
 CC -1- CATALYTIC ACTIVITY: Cleaves aggrecan at the 392-Glu-I-Ala-393  
 CC site.  
 CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular  
 CC matrix (by similarity).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, LUNG AND HEART. EXPRESSED  
 CC AT VERY LOW LEVEL IN PLACENTA AND SKELETAL MUSCLES.  
 CC -1- INDUCTION: BY INTERLEUKIN-1.  
 CC -1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT  
 CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.  
 CC -1- PTM: THE PRECURSOR IS CLEAVED BY A PURIN ENDOPEPTIDASE.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.  
 CC -1- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.  
 CC -1- CAUTION: HAS SOMETIMES BEEN REFERRED TO AS ADAMTS2.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: AB014588; BAA31663.1; -;  
 CC EMBL: AF148213; AAD41494.1; -;  
 CC EMBL: AY044847; AAL02262.1; -;  
 CC MEROFS: M12.221; -;  
 CC GeneW: HGNC:220; ADAMTS4.  
 DR MIM: 603876; -;  
 DR InterPro: IPR001762; Disintegrin.  
 DR InterPro: IPR001590; Reprolysin.  
 DR InterPro: IPR000884; TSP1.  
 DR InterPro: IPR000130; Zn\_M1peptidse.  
 DR Pfam: PF00090; tsp\_1; 1.  
 DR Pfam: PF01421; Reprolysin; 1.  
 DR SMART: SM00209; TSP1; 1.  
 DR PROSITE: PS00142; ZINC\_PROPEASE; 1.  
 DR PROSITE: PS50215; ADAM\_MERO; 1.  
 DR PROSITE: PS50092; TSP1; 1.  
 DR PROSITE: PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;

KW Extracellular matrix.  
 FT SIGNAL 1 51  
 FT PROPE 52 212  
 FT CHAIN 213 837  
 FT SITE 194 194  
 FT METAL 361 361  
 FT ACT\_SITE 362 362  
 FT METAL 365 365  
 FT METAL 371 371  
 FT METAL 437 519  
 FT DOMAIN 520 576  
 FT DOMAIN 577 685  
 FT DOMAIN 686 837  
 FT DOMAIN 247 252  
 FT CARBOHYD 68 68  
 FT CONFLICT 77 77  
 FT CONFLICT 626 626  
 FT CONFLICT 682 682  
 SQ SEQUENCE 837 AA; 90224 MW; 5DF9C9AC137DF41F CRC64;  
 Query Match 25.2%; Score 658.5; DB 1; Length 837;  
 Best Local Similarity 33.9%; Pred. No. 5.2e-41;  
 Matches 172; Conservative 63; Mismatches 156; Indels 117; Gaps 15;  
 QY 3 LGMASLLCAFLPLAANGPATPAODKAGOPTAAANOPRROGEEYQERAEPPGPH 62  
 DB 34 LVWLLILLASLP-----SRLASPLPREEIV-----PPE 65  
 QY 63 PLQARRRSGLVONIDOLYSGGKGYL--VYAGRRFLDLERDSVGIAGF----- 113  
 DB 66 KL-----NSVLPGSGAPARLLCRLOAFGETLLEEDSGVQVEGLVQYLGQ 114  
 QY 114 VPAGGTSAPMHRHRSCTFRGVDSRSLAYFDLCGG-LDGFPAVKHARITKPLLRP 172  
 DB 115 APELLGGAEP-----GYLTGTINDPESVASLHMDGGLLVLOYRGAELHLPLEG- 168  
 QY 173 WAEERGRVYGDGSARILHVYREGFSFALPPRASCEPP--ASTPEAH---EHAHPSN 227  
 DB 169 -----TPNAGGPGAILLRKSPA--- 187  
 QY 228 PGRAALASQLDQSLSPAGSGGPOTWMRRRRRSISRAROVELLVLVADASNARYGRL 287  
 DB 188 -SGCGPWC-----NWKALGSPSPRP--RRAKRFASLSFVETLVVADKMAAFHGAEL 238  
 QY 288 QHYLLTLASIANRLVSHAIEHRIILAVKYVYVLDKDKSLSEVSNNAATLKNFCKMHO 347  
 DB 239 KRLITVMAAAKAKRHPISIRNPVSLVYTRLVLSGEGGPVQGSAAOTLRSFCAMORG 298  
 QY 348 HNLGDDHEHYDAILFTREDLCGHSCDTGMADVGTICSPERSCAVIEDDGLHAFT 407  
 DB 299 LNTPESDPDHDTALFTPLRODLCVSTCDTGMADVGVCPDPAKALVEDDGLSATF 358  
 QY 408 VAHEIGHLLGHSDSKFCEETFGSTEDKR-LMSITLVSIDASKPWSKCTSATITEFLD 466  
 DB 359 AAHELGHVFMHLDNSKPCISLNGPLSTRHYMAPMAHVDEEPPWSPCSARFTDFLDN 418  
 QY 467 GHGNCLLDLPKROI-----LGDVYKD 488  
 DB 419 GYGHCLLDKPEAPLPTVTPPKDYDAD 446  
 RESULT 9  
 ATSG\_HUMAN STANDARD; PRT: 1629 AA.  
 ID ATSG\_HUMAN  
 AC Q9P2N4; Q9NR29;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE ADAMTS-9 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase  
 DE with thrombospondin motifs 9) (ADAM-TS 9) (ADAM-TS9).  
 GN ADAMTS9 OR KIA1312.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (SHORT ISOFORM).  
 RC TISSUE=Fetal;  
 RX MEDLINE=20396138; PubMed=10936055;  
 RA Clark M.E., Kellner G.S., Turbeville L.A., Boyer A., Arden K.A.,  
 RT Makl R.A.;  
 RT "ADAMTS 9, a novel member of the ADAM-TS/Metallopondin gene  
 family";  
 RL Genomics 67:343-350(2000).  
 RN [2]  
 RP SEQUENCE OF 159-1629 FROM N.A. (LONG ISOFORM).  
 RC TISSUE=Brain;  
 RX MEDLINE=20181126; PubMed=10718198;  
 RA Nagase T., Kikuno R., Ishikawa K.-I., Hirose M., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XVI.  
 RT The complete sequences of 150 new cDNA clones from brain which code  
 RT for large proteins in vitro.";  
 RL DNA Res. 7:65-73(2000).  
 CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular  
 CC matrix (By similarity).  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A  
 CC SHORT FORM. MAY BE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN ALL FETAL TISSUES.  
 CC EXPRESSED SLIGHTLY IN ADULT OVARY, PANCREAS, HEART, KIDNEY, LUNG,  
 CC PLACENTA. ALSO DETECTED IN SPINAL CORD AND BRAIN. NOT DETECTED IN  
 CC COLON, SMALL INTESTINE, TESTIS, LIVER, SKELETAL MUSCLE, SPLEEN OR  
 CC THYMUS.  
 CC -1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT  
 CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX (BY  
 CC SIMILARITY).  
 CC -1- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.  
 CC -1- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 11 TSP TYPE-1 DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF261918; AAF89106.1; -;  
 DR EMBL: AB037733; BAA92550.1; -;  
 DR HSSP: P15167; 1ATL.  
 DR MEROPS: M12.021; -;  
 DR Genew: HGNC:13202; ADAMTS9.  
 DR MIM: 605421; -;  
 DR InterPro: IPR001762; DisIntegrin.  
 DR InterPro: IPR002870; Pep\_M12B\_Proprep.  
 DR InterPro: IPR001590; Reptolysin.  
 DR InterPro: IPR000884; TSP1.  
 DR InterPro: IPR000130; Zn\_MTpeptidse.  
 DR Pfam: PF00090; tsp.1; 11.  
 DR Pfam: PF01421; Reptolysin. 1.  
 DR Pfam: PF01562; Pep\_M12B\_Proprep. 1.  
 DR SMART: SM00209; TSP1; 12.  
 DR PROSITE: PS50215; ADAM\_MERPRO; 1.  
 DR PROSITE: PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 DR PROSITE: PS50092; TSP1; 9.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 DR HydroLase: Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;  
 KM Repeat: Extracellular matrix; Alternative splicing.  
 FT SIGNAL 1 18  
 FT PROPEP 19 287 POTENTIAL.  
 FT CHAIN 288 1629 ADAMTS-9.  
 FT CHAIN 509 587 DISINTEGRIN-LIKE.  
 FT DOMAIN 589 642 TSP TYPE-1 1.

FT DOMAIN 645 752 CYS-RICH.  
 FT DOMAIN 753 880 SPACER.  
 FT DOMAIN 881 999 TSP TYPE-1 2.  
 FT DOMAIN 999 1053 TSP TYPE-1 3.  
 FT DOMAIN 1053 1108 TSP TYPE-1 4.  
 FT DOMAIN 1108 1156 TSP TYPE-1 5.  
 FT DOMAIN 1156 1239 TSP TYPE-1 6.  
 FT DOMAIN 1239 1295 TSP TYPE-1 7.  
 FT DOMAIN 1295 1332 TSP TYPE-1 8.  
 FT DOMAIN 1332 1386 TSP TYPE-1 9.  
 FT DOMAIN 1386 1445 TSP TYPE-1 10.  
 FT DOMAIN 1445 1554 TSP TYPE-1 11.  
 FT DOMAIN 1554 1562 TSP TYPE-1 11.  
 FT DOMAIN 1562 1632 POLY-SER.  
 FT DOMAIN 1632 1672 Cysteine switch (POTENTIAL).  
 FT SITE 223 223 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT SITE 223 223 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT ACT\_SITE 435 435 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT METAL 435 435 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT METAL 438 438 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT METAL 444 444 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 749 749 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 840 840 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1213 1213 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1267 1267 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARSPLIC 1064 1072 CLVCKGKH -> VRNCGCYFP (IN  
 FT VARSPLIC 1073 1629 SHORT ISOFORM).  
 FT VARSPLIC 1629 1629 MISSING (IN SHORT ISOFORM).  
 FT CONFLICT 367 367 P -> L (IN REF 1).  
 FT SEQUENCE 1629 AA; 182649 MW; C1C4CEFF58B941F CRC64;

Query Match 24.7%; Score 646; DB 1; Length 1629;  
 Best Local Similarity 32.3%; Pred. No. 16-39;  
 Matches 173; Conservative 72; Mismatches 195; Indels 96; Gaps 16;

3 LGASALLCAFLPLAAGPAAATPADDKAGQPTAAAOOPRRGEEVQ-----52  
 4 VSWATLLTLTLVR-DLAEWG-----SPDAAAVARKDLAPROYKLTLETGEYI 50  
 53 ---ERAEPGHPPL-AORRRSGLVONIDOLY-----SGGKYGVLVAGGRPL 99  
 51 VSRIRNALCEPPTVYHKKTRRSINSATDPMPARASSSSSTSOAHIRLSAFQOPL 110  
 100 LDLEKDSVGIAGFV-----PAGGCT--SAPMRHSCHPYRGTVDSPPSLAV 145  
 111 FNLTAN-----AGFIAPLFTVTLGTPGVNQTGFYSSEAEELKHCYKGVNNSHTAV 165  
 146 FDLGGDLGDFPAVKHARITLKLPL-LRGPMAREEKRRVGGDSARILHYTRGCFPEAL 203  
 166 ISLCSGMLGTFRSHDGDYFTEPLQSDMEQDEDEEOKKPH-----IYRRSAPQREPS 217  
 204 PRASCETPASPPEAHENHPAHS-----NPSGRALASOULDQSALSPG 248  
 218 TGNHACDT-----SEKNRHSKDKKTRARKWGERINLAGVVALNSGLATEASATG 270  
 249 GSGPOT---WMRRRRRSISRAROVELLVADASMARLYGRGLQHYLLTLASTANRLYSH 304  
 271 NKTNDTRERKTRHRTKRLPSYPRFEVLVADNRVSYNGENLOHYLTLTMSIVASTKYD 330  
 305 ASLENIIRLAUVKVVVVGSKDKSLEVSNAATTLKPFCKMOHNOHLDGDDHEHYAAIL 364  
 331 PSIGNLINIVYVLIHNEODGPSISFNAQTLLKFCOMOHKNSPGG---IHHDTAVL 387  
 365 FPREDDC-GHNSDPTGMADVGTICSPERSCAVIEDDGJLHAFTVAHEIGHLGLSHDS 423  
 388 LTRDDICRAHDKCDTLGLAEGLTICDPYRSCSISESGISTAFTHLAEHGVNMPHDDN 447  
 424 KCCEETFGSTEDKRLMSSILTSIDASKPWSKTSATITFEFLDGHGNCCLLDLPKQ 479  
 448 NKCKEE-GVKSQHVWAPLNFNTNPMWMSKCSKRYITFEFLDGYECCLNBEESR 502  
 RESULT 10



ATSA\_RAT STANDARD; PRT: 630 AA.  
 ID ATSA\_RAT  
 AC Q9ESP7: Q9ESP8: Q9ESP6:  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE ADAMTS-4 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 4) (ADAM-TS 4) (ADAM-TS4) (Aggrecanase 1) (Fragment).  
 GN ADAMTS4.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxId=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Wistar; TISSUE=Brain;  
 RX MEDLINE=20415831; PubMed=10961658;  
 RA Satoh K., Suzuki N., Yokota H.;  
 RT "ADAMTS-4 (a disintegrin and metalloproteinase with thrombospondin motifs) is transcriptionally induced in beta-amyloid treated rat astrocytes."  
 RL Neurosci. Lett. 289:177-180(2000).  
 CC -1- FUNCTION: CLEAVES AGGECAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE INVOLVED IN ITS TURNOVER. MAY PLAY AN IMPORTANT ROLE IN THE DESTRUCTION OF AGGECAN IN ARTHRITIC DISEASES.  
 CC -1- CATALYTIC ACTIVITY: Cleaves aggrecan at the 392-Glu-I-Ala-393 site.  
 CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (by similarity).  
 CC -1- TISSUE SPECIFICITY: ONLY EXPRESSED IN BRAIN.  
 CC -1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.  
 CC -1- PFM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.  
 CC -1- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: AB042272; BAB16474.1; -;  
 DR EMBL: AB042271; BAB16473.1; -;  
 DR EMBL: AB042273; BAB16475.1; -;  
 DR InterPro: IPR001762; Disintegrin.  
 DR InterPro: IPR001590; Reprolysin.  
 DR InterPro: IPR000884; TSP1.  
 DR InterPro: IPR000130; Zn\_MTPeptide.  
 DR Pfam: PF00090; tsp\_1; 2.  
 DR Pfam: PF01421; Reprolysin; 1.  
 DR SMART: SM00209; TSP1; 1.  
 DR PROSITE: PS02015; ADAM\_MPRO; 1.  
 DR PROSITE: PS00092; TSP1; 1.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 DR PROSITE: PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 KW Hydrolyase; Metalloprotease; zinc; Glycoprotein; Zymogen;  
 KW Extracellular matrix.  
 FT NON\_TER 1 1  
 FT PROPEP <1 5  
 FT CHAIN 6 630 BY SIMILARITY.  
 FT METAL 154 154 ADAMTS-4  
 FT ACT\_SITE 155 155 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT METAL 158 158 BY SIMILARITY.  
 FT METAL 164 164 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT METAL 164 164 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT DOMAIN 233 303 DISINTEGRIN-LIKE.  
 FT DOMAIN 316 367 TSP TYPE-1.  
 FT DOMAIN 368 478 CYS-RICH.

FT DOMAIN 479 630 SPACER.  
 FT DOMAIN 40 45 POLY-ALA.  
 FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 474 474 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SO SEQUENCE 630 AA; 68384 MW; 63442875316/CTEF CRC64;  
 Query Match 22.6%; Score 589.5; DB 1; Length 630;  
 Best Local Similarity 46.0%; Pred. No. 4.5e-36;  
 Matches 110; Conservative 45; Mismatches 77; Indels 7; Gaps 2;  
 QY 257 RRRRSISRARQVELLVADASMARLYGRGLQHYLLTLASTIANRLYSASTENHRIALAV 316  
 DB 1 RRTKFAISLRFVEVLVADDMKAFHAGGLKHYLLTVMMAAAAFKPSIRNPNVLVVT 60  
 QY 317 KVVVLGDKDKSLEYSKNAATLLKNCKWKQHOHNDLGDGHEHYAATLFTEDELCGHHSC 376  
 DB 61 RLVTILSGQEVPOGPPSAQOTLRSTWQKGLNPNDSDPHFDAILFTKODLCGVSTC 120  
 QY 377 DTLMADVGTICSPERSCAVIEDDGLAAFTVAHEIGHILGLSHDSKFCETFG-STED 435  
 DB 121 DALMGAGVGTCDPARSCAIVEDDGLSAFTAAHELGHVFMNLIDNSKPCANLNGGSSS 180  
 QY 436 KRMSIITSDASKPMKCTSATITTEFLDGHGCLLDLRKQI-----LGDYKRD 488  
 DB 181 RHVAPVAHVADPEEPWSPCSARFITDLDNGYGHCLDKPEAPLHPTEPGKDYDAD 239  
 RESULT 11  
 ATST7\_HUMAN STANDARD; PRT: 997 AA.  
 ID ATST7\_HUMAN  
 AC Q9UKP4;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE ADAMTS-7 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 7) (ADAM-TS 7) (ADAM-TS7).  
 GN ADAMTS7.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99395124; PubMed=10464288;  
 RA Hurskainen T.L., Hirohata S., Seldin M.F., Apte S.S.;  
 RT "ADAM-TS5, ADAM-TS6, and ADAM-TS7, Novel Members of a New Family of Zinc Metalloproteases".  
 RL J. Biol. Chem. 274:25555-25563(1999).  
 CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (by similarity).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN HEART, BRAIN, PLACENTA, LUNG, LIVER, SKELETAL MUSCLE, KIDNEY, PANCREAS.  
 CC -1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.  
 CC -1- PFM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.  
 CC -1- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 2 TSP TYPE-1 DOMAINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: AF140675; AAD56358.1; -;  
 DR HSSP: P15167; LATL.  
 DR MEROPS: M12.231; -;  
 DR Genew: HGNC:223; ADAMTS7.

MIM: 605009; -  
 DR InterPro: IPR001762; Disintegrin.  
 DR InterPro: IPR002870; Pep\_M12B\_propep.  
 DR InterPro: IPR001590; Reprolysin.  
 DR InterPro: IPR000884; TSPI.  
 DR InterPro: IPR000130; Zn\_MTPeptidase.  
 DR Pfam: PF00090; tsp\_1; 1.  
 DR Pfam: PF01421; Reprolysin; 1.  
 DR Pfam: PF01562; Pep\_M12B\_propep; 1.  
 DR SMART: SM00209; TSPI; 2.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 DR PROSITE: PS00215; ADAM\_MEPRO; 1.  
 DR PROSITE: PS0092; TSPI; 1.  
 DR PROSITE: PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 DR Hydrolyase: Metalloprotease; zinc; Signal; Glycoprotein; zymogen;  
 DR Repeat: Extracellular matrix.  
 KM SIGNAL  
 FT SIGNAL 1 27  
 FT PROPEP 23 232  
 FT CHAIN 233 937  
 FT SITE 204 204  
 FT METAL 388 388  
 FT ACT\_SITE 389 389  
 FT METAL 392 392  
 FT METAL 396 396  
 FT DOMAIN 462 537  
 FT DOMAIN 538 594  
 FT DOMAIN 595 697  
 FT DOMAIN 698 914  
 FT DOMAIN 915 990  
 FT CARBOHYD 94 94  
 FT CARBOHYD 693 693  
 FT CARBOHYD 778 778  
 SQ SEQUENCE 997 AA: 109694 MM: 6587044ED02FC104 CMC64;

Query Match 20.2% Score 527.5; DB 1; Length 997;  
 Best Local Similarity 31.5%; Pred. No. 3e-31; Mismatches 170; Indels 123; Gaps 21;  
 Matches 163; Conservative 62;

8 LLLCAFLPLAAGVAPATPAQDAKQO-----PPTAAAAO-----PRRQGEVQE 53  
 19 LLLCA-----LAPGAPPAAGRAALDIYVAVNDAGSGFLSYELMPALRAKDVSV 74  
 54 RAEPCHPPLAORRSKGLVONIDOLYSGGKVGVLVAGRRPLDLERDSVGIAGF 113  
 75 RRAAPAF-YELQYRGRELRFNLTAQO-----HLAPG--FVSETRRGGGLRAHI 121  
 114 VPAGGTSAPWRHRSN-----CFYRGTVDSGR--SLAVFDLGGDGFPAVKARYTLKP 167  
 122 -----RAITPACHLGEVQ-DPELEGGIAAISACDGLKGVQLSNEDEYFTEP 167  
 168 LIRGPAEEKGVYGDGSARILHYVTRGFSFEALPPRASCETPASTPEAHENPAHSN 227  
 168 LBSAP-----ARRGHAOPHYVKRQ-----APER----- 191  
 228 PSRRALASQLLDOSLSPAG-----GSGPQTV-----WRRRR-----RSISARQV 269  
 192 -----LAQRGSSAPSTCGVQVYVELLESRRERMOORWRPRLRLHQSSVSKEMV 244  
 270 ELLLVADASMARLYGR-GLQHYLLTLASIANRLYSHASIEHRIILAVVVVVLGDKDKSL 328  
 245 ETLVAVADAKMEVHGQPOVESYVLTIMNVAVGLFNDPSIGNPHITIVLVLLDEDEEL 304  
 329 EVSKNAATLTKNCKWOHONOLDGDHEDHYDAALFTREDLGC--HNSCDTLGMADVST 386  
 305 KITHHADNTLKSPCKWKQKSIKKGDANRPLHNDTALITLTKDLCAMNRCELTGLSHVAG 364  
 387 TCSPEKSCAVIEDDGLAAFTVAHEIGLLGLSHDSDS-KFCEETGEGSTDQR-LMSSTL 443  
 365 MCOPIHSCSINDDTGLPLAFYVAHEIGHSFGIOHDSGNDCEPV-----GKRPFIMSPOL 419  
 444 TSIDASKPSKCTSATITEFLDGHGNCCLLDIPROKIL 481  
 DB 420 LYDAAPLITWSKCSROYITTFILDRGWLCLDDPPAKDII 457

RESULT 12  
 ID AT12\_HUMAN STANDARD: PRT: 1593 AA.  
 AC P58397;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE ADAMTS-12 precursor (EC 3.4.24.-) (A disintegrin and metallopeptidase with thrombospondin motifs 12) (ADAM-TS 12) (ADAM-TS12).  
 GN ADAMTS12.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Fetal lung;  
 RA MEDLINE=21264577; PubMed=11279086;  
 RX Cal S.; Argueelles J.M.; Fernandez P.L.; Lopez-Otin C.;  
 RT "Identification, characterization, and intracellular processing of ADAM-TS12, a novel human disintegrin with a complex structural organization involving multiple thrombospondin-1 repeats.";  
 RL J. Biol. Chem. 276:17932-17940(2001).  
 CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).  
 CC -1- SURCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (by similarity).  
 CC -1- TISSUE SPECIFICITY: Expressed exclusively in fetal lung. Is widely expressed in gastric carcinomas and in cancer cells of diverse origin.  
 CC -1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE 1 DOMAINS ARE IMPORTANT FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX (BY SIMILARITY).  
 CC -1- PTM: THE PRECURSOR IS CLEAVED BY A PURIN ENDOPEPTIDASE.  
 CC -1- PTM: IS SUBJECTED TO AN INTRACELLULAR MATURATION PROCESS LEADING TO A FRAGMENT CONTAINING THE N-TERMINAL REGION INCLUDING THE METALLOPROTEINASE, DISINTEGRIN-LIKE, CYS-RICH AND TS-1 DOMAINS AND THE C-TERMINAL FRAGMENT CONTAINING THE SPACER 2 AND THE FOUR TS-1 DOMAINS.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.  
 CC -1- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 8 TSP TYPE-1 DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AJ250725; CAC20419.1; -  
 DR Genbank: HGNC:14605; ADAMTS12.  
 DR MIM: 606184; -  
 DR InterPro: IPR001762; Disintegrin.  
 DR InterPro: IPR002870; Pep\_M12B\_propep.  
 DR InterPro: IPR001590; Reprolysin.  
 DR InterPro: IPR000884; TSPI.  
 DR InterPro: IPR000130; Zn\_MTPeptidase.  
 DR Pfam: PF00090; tsp\_1; 6.  
 DR Pfam: PF01421; Reprolysin; 1.  
 DR Pfam: PF01562; Pep\_M12B\_propep; 1.  
 DR SMART: SM00209; TSPI; 8.  
 DR PROSITE: PS00215; ADAM\_MEPRO; 1.  
 DR PROSITE: PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 DR PROSITE: PS0092; TSPI; 2.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 KM Hydrolyase: Metalloprotease; zinc; Signal; Glycoprotein; zymogen;  
 KM Repeat: Extracellular matrix.  
 FT SIGNAL 1 25  
 FT PROPEP 26 240  
 BY SIMILARITY.

```

FT CHAIN 241 1593 ADAMTS-12.
FT DOMAIN 465 544 DISINTEGRIN-LIKE.
FT DOMAIN 545 596 TSP TYPE-1 1.
FT DOMAIN 597 700 CYS-RICH.
FT DOMAIN 701 826 SPACER 1.
FT DOMAIN 827 881 TSP TYPE-1 2.
FT DOMAIN 886 943 TSP TYPE-1 3.
FT DOMAIN 947 995 TSP TYPE-1 4.
FT DOMAIN 996 1315 SPACER 2.
FT DOMAIN 1316 1364 TSP TYPE-1 5.
FT DOMAIN 1367 1423 TSP TYPE-1 6.
FT DOMAIN 1426 1471 TSP TYPE-1 7.
FT DOMAIN 1426 1471 TSP TYPE-1 8.
FT DOMAIN 302 305 POLY-GLU.
FT SITE 208 208 CYSTEINE SWITCH (POTENTIAL).
FT METAL 392 392 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 393 393 BY SIMILARITY.
FT METAL 396 396 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 402 402 ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 215 215 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 485 485 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 685 685 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 790 790 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 951 951 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1104 1104 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1275 1275 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1300 1300 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1320 1320 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1371 1371 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1378 1378 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1503 1503 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1593 AA; 177545 MW; 07F9F48E63BD8A3 CRC64;

```

```

Query Match 19.1%, Score 500, DB 1: Length 1593;
Best Local Similarity 28.9%, Pred. No. 5.8e-29;
Matches 150; Conservative 72; Mismatches 185; Indels 112; Gaps 16;

```

```

OY 11 CAFRPLAAGVATPADKA---GQPTAAAOPTRRRQGEVQ-----ERAEP 57
DB 3 CAQRSMANLSVAQOLNFGALCYRQPOPGVRRPDRRQREFFITGLPEYHVGVRVDA 62
OY 58 PCH-----PHPLAQRRSKGLVONIDLY---SGGKGVYLVYAGRRFLDL-----ER 104
DB 63 SGHFLSTYGLHPTTSRRKRDLDGSEDMWYIRISHEEKDLFNLVNGFLSNSTIMEKR 122
OY 105 DGSVGIAGFVAGGCTSPWRHRSCHCYRGTY--DGSPRSLAVFDLGGDGFVAVKANR 162
DB 123 YGNLSHVKKM-----ASSAPLCHLS-----GVLQOGTRVGTAAALSACHGLTGFOLPHGD 173
OY 163 YTLKLLRGPMAEKKGKGVYGDGSAKILHVTYREGFSFPAUPRRASCPTATPEHEHA 222
DB 174 FFIIEVKKHPLVE-----GTYHPIHYVRKQ-----KVPEKKE-- 205
OY 223 PAHNSPGRALASQLLDOSALSPAGG-----SGPQTMWRR-----RRSI 263
DB 243 SKERNVEFLVADTMKIEHGESENVESYILITMNVGTGLFNHPSIGMAIHIVVRLILLE 302
OY 323 DKDKSLSEYKNAATLKNPKQWQHONOLGDDEEHYDAALIFTEDEDCG--HSGCDTIG 380
DB 303 EEEQGLKLYHNAEKLTSFCKMOKSINPKSDLPVHDDVAILLRKIDCAGNRPETIG 362
OY 361 MADVGTICSPERSCAVIEDDGLAAFTVAHEIGHLGLSHDSDSKCEETFGSTEDKRLMS 440
DB 363 LSHLSKMGCPHNSCINNESGLPLAFTLAHEIGHSFGIOHDKENDCEPVG--RHPIYLS 420
OY 441 SLTSTIDASKPWSKCTSATITFEFLDGGHNCCLLDLPKRO 479

```

```

DB 421 ROLQYDPTLTWSKCSBEYITRFLDRGWFCLUDDIPKRR 459
RESULT 13
ID AT10_HUMAN STANDARD; PRT; 1077 AA.
AC Q9H324;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE ADAMTS-10 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 10) (ADAM-TS10) (ADAM-TS10) (Fragment).
GN ADAMTS10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
RX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Apte S.S.;
RT "ADAM-TS10: a novel member of the ADAM-TS family containing multiple
RL thrombospondin type I repeats.";
CC Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (by similarity).
CC -1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE 1 DOMAINS ARE IMPORTANT
CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC -1- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 5 TSP TYPE-1 DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF163762; AAC35563.1; -.
DR MEROPS: M12.235; -.
DR Genew: HGNC:13201; ADAMTS10.
DR InterPro: IPR001762; Disintegrin.
DR InterPro: IPR002870; Rep_M12B_propep.
DR InterPro: IPR001590; Reptolysin.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00090; TSP1; 5.
DR Pfam: PF00130; Z0_MTPeptide.
DR Pfam: PF01421; Reptolysin; 1.
DR Pfam: PF01562; Rep_M12B_propep; 1.
DR SMART: SM00209; TSP1; 5.
DR PROSITE: PS50215; ADAM_MERO; 1.
DR PROSITE: PS50092; TSP1; 2.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
DR PROSITE: PS00427; DISINTEGRIN_1; FALSE NEG.
KW Hydrolase; Metalloprotease; Zinc; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix.
FT NON_TER 1
FT PROPEP <1 207 BY SIMILARITY.
FT CHAIN 208 1077 ADAMTS-10.
FT METAL 366 366 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 367 367 BY SIMILARITY.
FT METAL 370 370 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 376 376 ZINC (CATALYTIC) (BY SIMILARITY).
FT DOMAIN 434 520 ZINC (CATALYTIC) (BY SIMILARITY).
FT DOMAIN 578 679 DISINTEGRIN-LIKE.
FT DOMAIN 680 802 CYS-RICH.
FT DOMAIN 521 577 SPACER.
FT DOMAIN 577 799 TSP TYPE-1 1.
FT DOMAIN 799 860 TSP TYPE-1 2.
FT DOMAIN 862 918 TSP TYPE-1 3.
FT DOMAIN 922 976 TSP TYPE-1 4.

```

```

FT DOMAIN 981 1031 TSP TYPE-1 5.
FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 297 297 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 714 714 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 769 769 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 866 866 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1077 AA; 118072 MW; 391ADE18DCBFB587 CRC64;

Query Match
Best Local Similarity 28.9%; Pred. No. 2,7e+28;
Matches 138; Conservative 68; Mismatches 168; Indels 104; Gaps 16;

QY 61 PHLAQRNRKSLGVNIDLYSGGKGYLYVAGRRFLDLERDGSVGIAGFVPGGT 120
DB 32 PPPRRRRRGTCGTA-----TAESRLFYKVASPTHTFLNLTNRSSRL-LAGHVSVEYWT 82
QY 121 --SAPRRH--RSHCFYRGTVDSPRSLAV-FDLGGLDGFPAVKHARYLKPLLRP--- 172
DB 83 REGLAQRARPHCLYAGHLGQASSSHAISTCGGLHGLIYADEEYLTLEPLHGPKGS 142
QY 173 MAEEKGR--VYGDGSRILHYTRGCFEALPPRASCETPASTPEAHEHAPHSNPSG 230
DB 143 RRPESGRHYVKKRSSLRPHLDTAG-----VDEKPMKGRP-- 180
QY 231 RAALASQLLDQSALSPAGSGPOTWRRR-----RSTISARQVE 270
DB 181 -----WMLRTLKPPAPPLCNETERGCPGLKRSVSRERYE 216
QY 271 LLLVADAMARLYG-RGLQHYLLTLASINRLYSHASINHLRLAVKVVVGLDCKSLE 329
DB 217 TLVVADKMMVYHGRDVEQYVLAIMNIYAKLPDSSLSVNIITLRLILTEDPTE 276
QY 330 VKNMATTLKNECKMOH-----OHNLQDDHEEHYDAIIFTRFEDLCGHS--CDTLG 380
DB 277 ITNHGKSLDSCCKMKQKSIYVNSHGNALPENGVAHNOVAVLITRDICITKKNPCGTIG 336
QY 381 MADVGTICPESCAVETDGGHAAFTVAHEIGHLGLSHD-SKFCETPSTEDKRLM 439
DB 337 LAPVGMCMERSCSVNEDIGLPQAFITIAHEIGHFGMHDGVCNSGAR--GQDPAXLM 394
QY 440 SSILTSIDASKPMSKTSATITEFLDDGHCNLLDPRQO-----ILGGDYKDDDDK 491
DB 395 AAHTTKTPFVWSSCNBDYITSLDGLGLCLNNRPQDVEYVPIVAPQAVDADEQ 452

RESULT 14
ATTS5_BOVIN
ID ATTS5_BOVIN STANDARD: PRT: 207 AA.
AC 09T792;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 15-JUN-2002 (Rel. 41; Last annotation update)
DE ADAMTS-5 (EC 3.4.24.-) (A disintegrin and metalloproteinase with
thrombospondin motifs 5) (ADAM-TS 5) (ADAM-TS5) (Aggrecanase-2) (ADMP-
2) (ADAM-TS 11) (Fragment).
DE GN
GN ADAMTS5.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_Taxid:9913;
RN NCBI (1)
RN SEQUENCE FROM N.A.
RX MEDLINE=99333677; PubMed=10403768;
RA Flannery C.R., Little C.B., Hughes C.E., Caterson B.;
RT "Expression of ADAMTS homologues in articular cartilage.";
RL Biochem. Biophys. Res. Commun. 260:318-322(1999).
RN (2)
RN SEQUENCE FROM N.A.
RX MEDLINE=20092827; PubMed=10625599;
RA Curtis C.L., Hughes C.E., Flannery C.R., Little C.B., Harwood J.L.,
RA Caterson B.;
```

```

RT "n-3 fatty acids specifically modulate catabolic factors involved in
articular cartilage degradation.";
RL J. Biol. Chem. 275:721-724(2000).
CC -1- FUNCTION: CLEAVES AGGREGAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE
INVOLVED IN ITS TURNOVER. MAY PLAY AN IMPORTANT ROLE IN THE
DESTRUCTION OF AGGREGAN IN ARTHRITIC DISEASES. MAY PLAY A ROLE IN
PROTEOLYTIC PROCESSING MOSTLY DURING THE PERI-IMPLANTATION PERIOD.
CC -1- CATALYTIC ACTIVITY: Cleaves aggrecan at the 392-Glu-I-Ala-393
site.
CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
matrix (By similarity).
CC -1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC -1- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY
SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC -1- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS AT LEAST 1 TSP TYPE-1 DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF192771; AAF07177.1; -
DR InterPro: IPR001590; Reprolysin.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR000130; Zn_MTPeptidase.
DR Pfam: PF00090; tsp_1; 1.
DR Pfam: PF01421; Reprolysin; 1.
DR SMART: SM00209; TSP1; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
DR PROSITE: PS50215; ADAM_MEPRO; 1.
DR PROSITE: PS50092; TSP1; 1.
KW Hydrolase; Metalloproteinase; Zinc; Glycoprotein; Extracellular matrix.
FT NON_TER 1 1
FT METAL 8 8
FT ACT_SITE 9 9
FT METAL 12 12
FT METAL 18 18
FT DOMAIN 83 164
FT DOMAIN 165 207
FT CARBOHYD 96 96
FT NON_TER 207 207
SQ SEQUENCE 207 AA; 22575 MW; E648BEA73A3F86EE CRC64;

Query Match
Best Local Similarity 15.8%; Score 413; DB 1; Length 207;
Matches 78; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 403 HAAFTVAHEIGHLGLSHDSKFCETPSTEDKRLMSILTSIDASKPMSKTSATITE 462
DB 1 HAAFTVAHEIGHLGLSHDSKFCETPSTEDKRLMSILTSIDASKPMSKTSATITE 60
QY 463 FLDDGHCNCLDLPRKOLIG 482
DB 61 FLDDGHCNCLDLPRKOLIG 80

RESULT 15
ATTS6_HUMAN
ID ATTS6_HUMAN STANDARD: PRT: 860 AA.
AC 09UKP5;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 15-JUN-2002 (Rel. 41; Last annotation update)
DE ADAMTS-6 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
with thrombospondin motifs 6) (ADAM-TS 6) (ADAM-TS6).
GN ADAMTS6.
```



**THIS PAGE BLANK (USPTO)**